В МИРЕ НАУКИ

SCIENTIFIC AMERICAN

Издание на русском языке



Апрель **4** *1983*

СИНТЕТИЧЕСКИЕ ВАКЦИНЫ

Издательство **МИР** *предлагает:*

Я. Мусил, О. Новакова, К. Кунц

СОВРЕМЕННАЯ БИОХИМИЯ В СХЕМАХ

Перевод с английского

2-е издание



Книга, написанная группой чешских авторов, является совершенно необычным уникальным изданием по биохимии. В последние годы на русском языке появился ряд прекрасных книг по биохимии, предназначенных главным образом для специалистов, работающих в этой области. Данное издание принципиально отличается от них по крайней мере двумя особенностями. Во-первых, своим объемом. Фундаментальные основы биохимии и новейшие достижения этой науки изложены в ней очень лаконично. Концентрация огромного фактического материала достигнута за счет того, что центральное место отведено очень наглядным цветным схемам, а краткий текст лишь поясняет их и включает определение основных понятий. Это талантливо составленный своего рода конспект различных разделов биохимии.

Вторая особенность книги заключается в том, что она рассчитана на самый широкий круг читателей. Она окажется полезной для всех, кто изучает биохимию как наиболее краткая форма изложения, и особенно для тех, кто не имея специальной подготовки, столкнулся с ее проблемами. Таким образом, она будучи строго научным изданием максимально облегчает знакомство с этой областью науки.

Русское издание «Современной биохимии в схемах» значительно более полное по сравнению с оринигалом, поскольку авторы включили в него шесть новых глав, посвященных вопросам, которые ранее в книге не рассматривались.

Многочисленные положительные отзывы читателей на 1-е издание позволяют с уверенностью сказать, что книга будет широко использоваться исследователями, работающими в смежных с биохимией областях, студентами и аспирантами, изучающими биохимию, а также преподавателями школ, техникумов и педагогических вузов.

1984, 26 изд. л. Цена 2 руб.

В МИРЕ НАУКИ

Scientific American · Издание на русском языке

ИЛЛЮСТРИРОВАННЫЙ ЖУРНАЛ

ПЕРЕВОД С АНГЛИЙСКОГО · ВЫХОДИТ 12 РАЗ В ГОД · ИЗДАЕТСЯ С 1983 ГОДА

ИЗДАТЕЛЬСТВО «МИР» МОСКВА

№ 4 · АПРЕЛЬ 1983

В номере:

SCIENTIFIC AMERICAN

BOARD OF EDITORS
Gerard Piel
PUBLISHER
Dennis Flanagan
EDITOR
Brian P. Hayes
ASSOCIATE EDITOR
Philip Morrison
BOOK EDITOR
Francis Bello
John M. Benditt
Peter G. Brown
Michael Feirtag
Jonathan B. Piel
John Purcell
James T. Rogers
Armand Schwab, Jr.
Joseph Wisnovsky

Samuel L. Howard
ART DIRECTOR
Richard Sasso
PRODUCTION MANAGER
George S. Conn
GENERAL MANAGER

В МИРЕ НАУКИ

ГЛАВНЫЙ РЕДАКТОР
С. П. Капица
ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА
Л. В. Шепелева
НАУЧНЫЕ РЕДАКТОРЫ
З. Е. Кожанова
О. К. Кудрявов
Т. А. Румянцева
А. М. Смотров
ЛИТЕРАТУРНЫЕ РЕДАКТОРЫ
Л. И. Леонова
М. М. Попова
ХУДОЖЕСТВЕННЫЙ РЕДАКТОР
С. А. Стулов
ТЕХНИЧЕСКИЙ РЕДАКТОР
Л. П. Чуркина
КОРРЕКТОР
И. И. Дериколенко
ОФОРМЛЕНИЕ ОБЛОЖКИ, ТИПОГРАФИКА РУССКОГО ИЗДАНИЯ, МАКЕТ СМЕННЫХ ПОЛОС:
М. Г. Жуков
ТИТУЛЬНАЯ НАДПИСЬ, ШРИФТОВЫЕ РАБОТЫ:
В. В. Ефимов

АДРЕС РЕДАКЦИИ
129820, Москва, И-110, ГСП
1-й Рижский пер., 2
ТЕЛЕФОН РЕДАКЦИИ
286.2588

На обложке



СИНТЕТИЧЕСКИЕ ВАКЦИНЫ

Структура полиэдрической белковой оболочки, или капсида, вируса карликовой кустистости томатов (изображена синим цветом). Капсид состоит из 180 молекул белка; каждая из 60 треугольных граней капсида образована тремя такими молекулами; 12 поверхностных доменов белка показаны оранжевым цветом. Рисунок сделан по изображению, полученному А. Олсоном с помощью ЭВМ. Анализ таких изображений — один из этапов разработки синтетических вакцин (см. статью Р. Лернера «Синтетические вакцины», с. 4). Обычно вакцины против возбудителей вирусных инфекций содержат целые вирусные частицы, синтетическая же вакцина — это искусственно синтезированный пептид, т.е. короткий отрезок полипептидной цепи, который имитирует участок поверхностного белка вируса. При разработке синтетической вакцины необходимо знать, с какими участками молекулы вирусного белка связываются вырабатываемые против него антитела. Исследования показали, что антитела реагируют практически с любым участком на поверхности белка, если он не погружен в глубь вирусной частицы и доступен молекулам, находящимся в кровотоке. Т. О'Доннелл, М. Коннолли и А. Олсон разработали методы построения графиков, с помощью которых можно убедиться, что данный пептид действительно имитирует участок поверхностного вирусного белка и поэтому должен вызывать образование антител против целых вирусных частиц. Уже получены синтетические вакцины против вирусов гриппа, ящура, гепатита и против дифтерийного токсина. В будущем подобные препараты позволят эффективно бороться с опасными заболеваниями.

Иллюстрации

ОБЛОЖКА: Gabor Kiss

СТР.	АВТОР/ИСТОЧНИК
50—57	Andrew Tomko
37—47	Ilil Arbel
5	Arthur J. Olson, Research Institute of Scripps Clinic
6—7	Patricia J. Wynne
8—11	Arthur J. Olson, Research Institute of Scripps Clinic
12	Patricia J. Wynne
16—23	Walken Graphics
27	R. F. Bonifield, с любезного разрешения Granger Collection
28—34	George V. Kelvin
61	High Altitude Observatory
62	Jay M. Pasachoff, Williams College (*вверху*) High Altitude Observatory (*внизу*)
63	High Altitude Observatory
64	American Science and Engineering, Inc.
65	Goddard Space Flight Center (*вверху*), High Altitude Observatory (*внизу*)
66	Naval Research Laboratory
67—69	Allen Beechel
72	David M. Dennis
73	Patricia J. Wynne
74	Charles W. Myers, American Museum of Natural History
75	Lois W. Tice, National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases
76	Patricia J. Wynne
78	Enid Kotschnig
80	John Moss, Black Star
82—85	Alan D. Iselin
86	Walter C. Leedy, Jr., Cleveland State University (*вверху*) Alan D. Iselin (*внизу*)
87	Alan D. Iselin
88	John Moss, Black Star
90—94	Michael Goodman

Об авторах

Sandra S. Batie и Robert G. Healy (САНДРА С. БЕЙТИ и РОБЕРТ Г. ХИЛИ «Будущее сельского хозяйства США») — экономисты, особенно интересуются проблемами природных ресурсов и землеустройства. С. Бейти окончила Вашингтонский университет (г. Сиэтл, шт. Вашингтон), а степени магистра и доктора получила в Университете шт. Орегон. С 1973 г. она доцент кафедры экономики сельского хозяйства в Вирджинском политехническом университете. Р. Хили получил ученые степени бакалавра, магистра и доктора * в области экономики в Калифорнийском университете в Лос-Анджелесе. С 1975 г. он штатный сотрудник Фонда охраны природных ресурсов в Вашингтоне. В их статье использованы данные широкого комплексного исследования, полные результаты которого приведены в книге «Будущее американского сельского хозяйства как стратегического ресурса», написанной группой авторов с участием С. Бейти и Р. Хили.

Claudio Rebbi (КЛАУДИО РЕББИ «Решеточная теория удержания кварков») с 1977 г. работает в Брукхейвенской национальной лаборатории, специалист в области теории элементарных частиц. К. Ребби изучал физику в Туринском университете, где в 1967 г. получил докторскую степень. Затем стажировался в Калифорнийском технологическом институте, преподавал в Университете г. Триест и вел научно-исследовательскую работу в Европейской организации ядерных исследований (ЦЕРН) в Женеве. В 1974 г. К. Ребби был приглашен в Массачусетский технологический институт на должность доцента. В течение 1980-1981 гг. он снова занимался исследовательской работой в ЦЕРНе.

Richard A. Lerner (РИЧАРД А. ЛЕРНЕР «Синтетические вакцины») заведует отделом молекулярной биологии в научно-исследовательском институте при клинике Скриппса в г. Ла-Холья (шт. Калифорния). Степень доктора медицины он получил в 1964 г. в Станфордском университете, после чего год работал в ординатуре Медицинской школы Станфордского университета. С тех пор Лернер постоянно связан с клиникой Скриппса, если не считать 1968-1970 гг., когда он работал в Филадельфии в Вистаровском институте анатомии и биологии.

Eitan Abraham, Colin T. Seaton and S. Desmond Smith (АЙТЕН ЭЙБРЭХЭМ, КОЛИН Т.СИТОН и С.ДЕЗМОН СМИТ «Оптический компьютер») работают в Университете Эрио-Уатта в Эдинбурге. А. Эйбрэхэм — старший научный сотрудник кафедры физики. Родился в Израиле, в 1975 г. окончил Университет в Буэнос-Айресе, в 1979 г. получил докторскую степень в Манчестерском университете. К. Ситон в 1979 г. окончил Университет Эрио-Уатта, где сейчас работает над докторской диссертацией. Его научная работа связана с изучением нелинейных оптических эффектов в полупроводниках и созданием оптических устройств обработки сигналов. С.Д. Смит — профессор физики Университета Эрио-Уатта и одновременно декан факультета естественных наук. Окончил Бристольский университет, получил степени доктора философии (1956 г.) в Университете Ридинга и доктора наук (1966 г.) в Бристольском университете. С.Д. Смит является членом секции физики Королевского научного общества в Эдинбурге, а также членом Лондонского Королевского общества **

Jeremy M. Wolfe (ДЖЕРЕМИ М.ВУЛФ «Скрытые зрительные процессы») читает вводный курс по психологии и ведет научную работу в Массачусетском технологическом институте (МТИ). Он родился в Лондоне, в 1977 г. окончил с отличием Принстонский университет. Темой его диссертации, над которой он работал в МТИ, было изучение бинокулярного зрения; за эту работу в 1981 г. ему была присуждена степень доктора. Вулф пишет о себе: «В школе я думал, что стану юристом или священнослужителем. Но как-то во время летних каникул я подрабатывал в научно-исследовательском центре при фирме Bell Laboratories и там меня обратили в служителя науки — я увлекся экспериментальной психологией зрительных процессов. Я часами просиживал в темной комнате, наблюдая за едва различимыми вспышками света разных цветов, после чего преисполнился намерением заставить других людей делать то же самое». Свободное время Дж. Вулф любит проводить у себя дома в Кембридже (шт. Массачусетс) «с женой, двумя кошками и двухметровым бирманским питоном».

Richard Wolfson (РИЧАРД ВОЛЬФСОН «Активная солнечная корона») — доцент кафедры физики в Колледже Мидлбери. Учился в Массачусетском технологическом институте, а затем в Суортморском колледже, который закончил в 1969 г., получив степени бакалавра по физике и по философии. Позже Р. Вольфсон получил в Мичиганском университете степень магистра в области наук об окружающей среде. Затем в Дортмутском колледже получил степень доктора физических наук. По его словам, его интересуют исследования поведения газов и плазмы в космосе, в том числе аккреция вещества в источниках космического рентгеновского излучения, солнечный и звездный ветер, а также солнечная корона. «В 1980-1981 гг. меня пригласили работать на высокогорную обсерваторию Национального центра по исследованию атмосферы; там я занялся теоретическими вопросами динамики солнечной короны. Мне хотелось бы поблагодарить сотрудников обсерватории за помощь в работе и всех тех, кто предоставил мне фотографии солнечной короны».

Charles W. Myers, John W. Daly (ЧАРЛЗ У. МАЙЕРС и ДЖОН У. ДЭЛИ «Ядовитые лягушки») работают вместе с 1966 г. Они сочетают работу в экспедициях с лабораторными исследованиями. Майерс и Дэли — сторонники комплексного подхода к изучению живых существ. Ядовитых лягушек они изучают с позиций как общей биологии, так и биохимии. Ч.У. Майерс заведует Отделом герпетологии в Американском музее естественной истории. В 1960 г. он окончил Флоридский университет со степенью бакалавра. Степень магистра получил в 1962 г. в Университете Южного Иллинойса. В 1964-1967 гг.работал в качестве приглашенного научного сотрудника в лаборатории им. Горгаса (Панама). В Музей естественной истории его приняли в 1968 г. В 1970 г. Ч.У. Майерс получил степень доктора в Канзасском университете. Дж. У. Дэли — руководитель лаборатории биоорганической химии в Национальном институте артрита, диабета, заболеваний желудочно-кишечного тракта и почек. Дэли имеет три ученые степени в области химии — степени бакалавра (1954 г.) и магистра (1955 г.) он получил в Колледже шт. Орегон, а степень доктора (1958 г.) — в Станфордском университете.

Walter C. Leedy, Jr. (УОЛТЕР К. ЛИДИ, мл. «Веерные своды») — историк архитектуры средних веков. Работает в Кливлендском университете. Учился в Мичиганском университете и закончил там же аспирантуру. В 1966 — 1967 гг. был стажером-стипендиатом Фонда Костюшко в ПНР. В 1972 г. Институт искусств им. Куртолда Лондонского университета присудил ему степень доктора в области истории искусств.

* Степень *бакалавра* соответствует высшему образованию; степень *магистра* — нашему уровню аспирантуры без защиты диссертации. Низшая ученая степень — *доктора философии* — званию кандидата наук. Степень *доктора* в какой-либо науке — званию доктора наук. — *Прим. ред.*

** Британская академия наук. — *Прим. перев.*

Синтетические вакцины

Удалось синтезировать короткие цепочки аминокислот, которые имитируют определенные участки поверхности вирусного белка. С их помощью можно индуцировать образование антител с заранее известной специфичностью и создать иммунитет против вируса

РИЧАРД А. ЛЕРНЕР

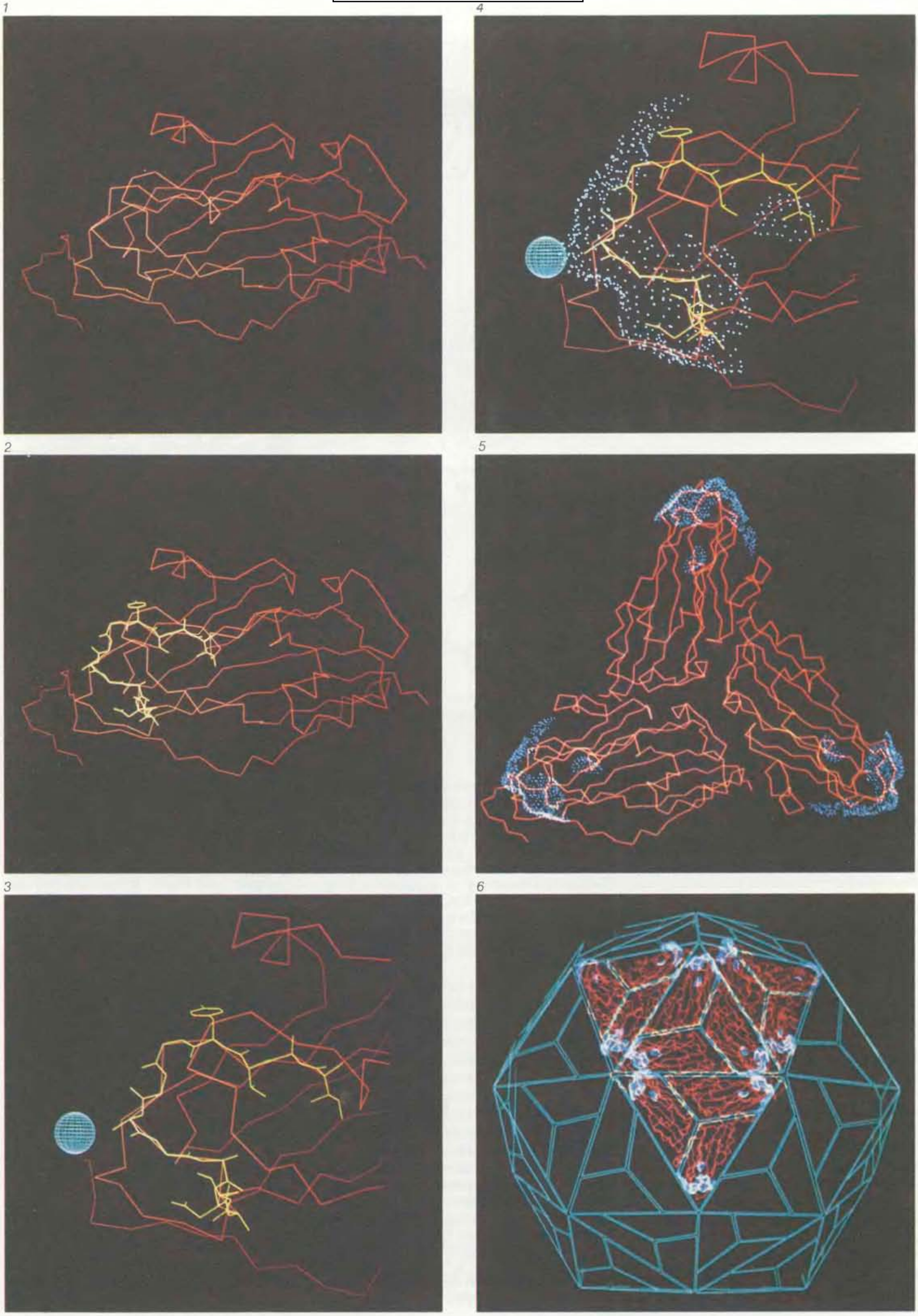
ВАКЦИНАЦИЯ — одно из крупнейших достижений биологии. С ее помощью удалось победить оспу и взять под контроль такие заболевания, как дифтерия, полиомиелит и корь. Однако техника вакцинации до сих пор несовершенна; самое главное, невозможно гарантировать абсолютную безопасность вакцины. Иммунизация против вирусной инфекции достигается тем, что в организм человека или животного вводят вирус; в ответ на это образуются специфические антитела, которые обезвреживают возбудителя болезни. Вирус должен быть предварительно ослаблен или убит, чтобы он не мог вызвать заболевание. Ослабленным вирус становится после пассажей через организм какого-либо животного, во время которых он мутирует и становится менее вирулентным для человека. (Если человеку ввести вирус коровьей оспы, то у него возникает иммунитет против натуральной оспы, т.е. вирус коровьей оспы — это ослабленный вирус, получившийся естественным путем.) Однако ослабленный вирус еще жив и сохраняет способность изменяться. Он может снова мутировать так, что его вирулентность для человека возрастет. Даже если вирус не мутирует, он может обладать каким-либо неизвестным медленно проявляющимся свойством вроде тех, что наблюдаются у так называемых латентных вирусов. Неприятных последствий можно избежать, если убить и таким образом инактивировать вирус, однако случалось, что вирус инактивировался неполностью и вызывал заболевание.

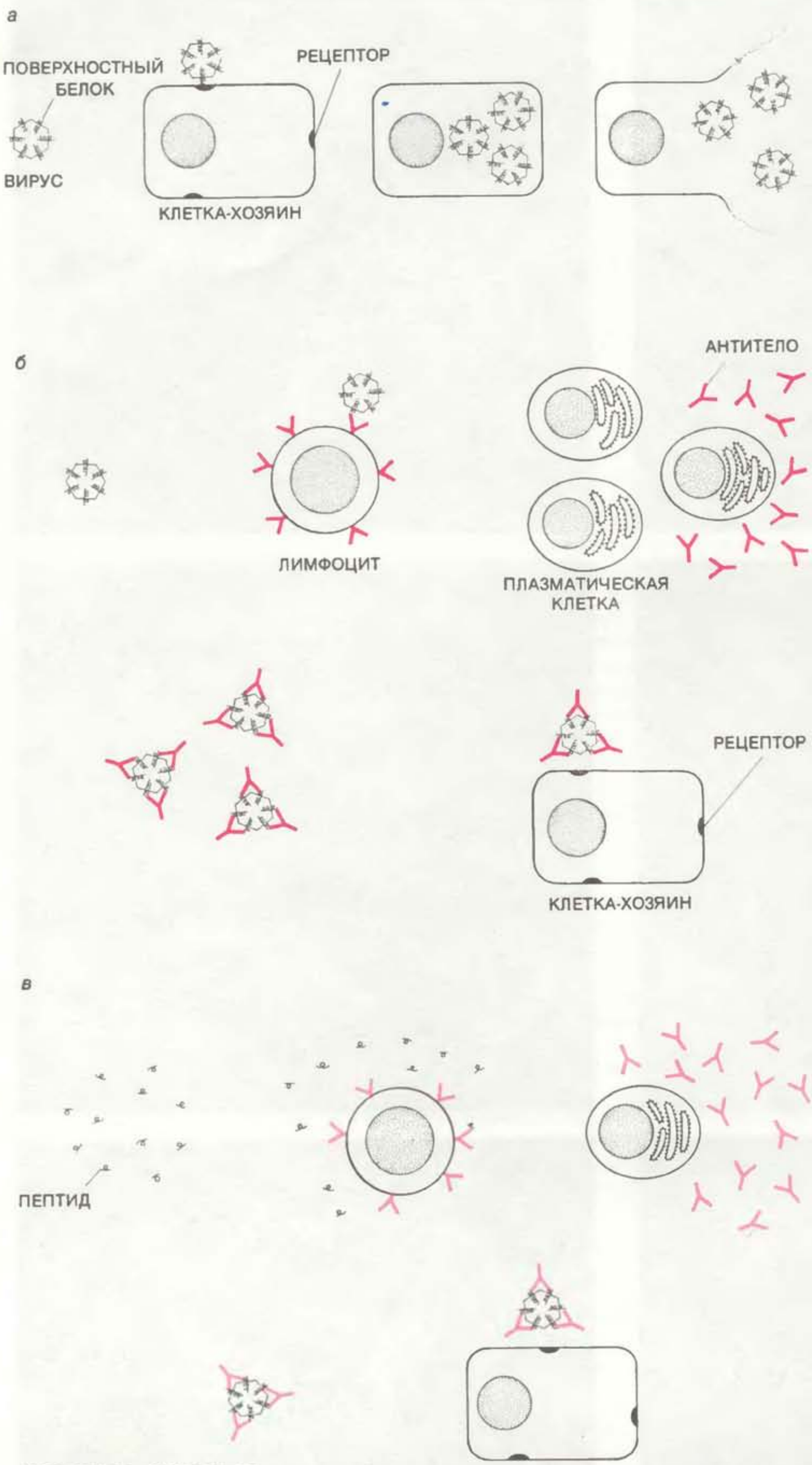
Другой круг проблем относится и к вакцинам с убитыми вирусами, и к вакцинам с ослабленными вирусами. Все устройства типа ферментеров, в которых выращивают вирусы, представляют собой потенциальный источник распространения возбудителя заболевания. Кроме того, поскольку вирусы размножаются только в живых системах, их обычно выращивают в культуре клеток или в куриных эмбрионах либо выделяют из крови зараженных животных (в случае вируса гепатита В — из крови больных людей); в культуре клеток, в куриных эмбрионах и в крови могут быть трудно обнаружимые агенты — в частности другие вирусы, — которые загрязнят вакцину. Многие вакцины даже с убитыми вирусами необходимо хранить при пониженной температуре, а создать надежные системы для этого нелегко.

В связи с этими трудностями возник большой интерес к синтетическим вакцинам, т.е. к таким вакцинам, которые содержат не целые вирусы, а искусственно синтезированные короткие пептиды, имитирующие небольшие участки внешней оболочки вируса и способные вызывать образование антител против этого вируса. В ряде лабораторий, в том числе и в нашей (Научно-исследовательский институт клиники Скриппса), уже получены первые экспериментальные вакцины такого рода. Выяснилось много нового об иммунологически важных особенностях структуры белков. Умение синтезировать пептиды, вызывающие у животных образование антител с вполне определенной и заранее известной специфичностью, дает молекулярным биологам мощное средство исследования; пожалуй, это не менее важно, чем применение синтетических вакцин в медицине.

ПРЕЖДЕ ЧЕМ рассказать о результатах наших экспериментов и новых данных по структуре белков, мне бы хотелось остановиться на том, каким образом иммунная система защищает организм от вирусной инфекции. Любой вирус содержит небольшой объем генетической информации, записанной в ДНК или РНК. Нуклеиновая кислота вируса окружена оболочкой (капсидом), состоящей из многих молекул одного или нескольких видов белка. Капсид в свою очередь может быть окружен мембраной, усеянной белками. Инфекция начинается с того, что специфический участок на поверхности вируса связывается со специфическим рецептором на поверхности клетки. Вирус проникает сквозь мембрану в клетку и вынуждает ее биосинтетический аппарат образовывать копии вируса. Новые вирусные частицы выходят наружу и заражают другие клетки. Иммунная система организма может блокировать инфекционный процесс, если животное сталкивалось ранее с этим же вирусом или было иммунизировано вакциной против него. В обоих случаях специальные клетки иммунной системы — лимфоциты — окажутся запрограммированными для узнавания вируса. Как только поверхностный рецептор одного из таких лимфоцитов вступает в контакт с вирусом, запуска-

СИНТЕТИЧЕСКИЕ ВАКЦИНЫ конструируют с помощью полученных на ЭВМ графиков. Рисунки изображений (получены А. Олсоном из Научно-исследовательского института клиники Скриппса) демонстрируют, каким образом можно определить, какие участки молекулы вирусного белка находятся на поверхности вирусной частицы и поэтому доступны для антител. *1.* Остов поверхностного домена белка внешней оболочки вируса кустистой карликовости томатов (координаты определили С. Харрисон и его коллеги, Гарвардский университет). *2.* Отдельный пептид белка (*желтый*); в боковых цепях аминокислотных остатков отмечены детали на атомном уровне. *3.* Тот же пептид в увеличенном виде; слева показана сфера, представляющая молекулу воды. *4.* Сфера «катится» по поверхности пептида, и получается карта участков поверхности, доступных молекулам воды. Карта строится по алгоритму, разработанному М. Коннолли: в тех местах, где соотношение вандерваальсовых радиусов сферической модели молекулы воды и атомов пептида не исключает тесного контакта молекулы воды с атомами пептида, ставится точка (учитывается и влияние атомов остальной части молекулы белка). Такая точечная карта показана для комплекса из трех молекул белка на поверхности вирусной частицы (*5*) и для четырех (из 60) таких комплексов, покрывающих вирусную частицу (*6*).


1
2
3
4
5
6



ВИРУСНУЮ ИНФЕКЦИЮ можно предотвратить, введя в организм человека или животного вакцину. *а* — Вирус присоединяется к рецептору на поверхности клетки, проникает внутрь и заставляет клетку синтезировать копии вирусной частицы; клетка гибнет, новые вирусы выходят наружу. *б* — Вакциной может служить ослабленный или убитый вирус. Связывание такого вируса с рецептором стимулирует лимфоцит к делению, и возникает клон плазматических клеток, секретирующих антитела. Когда такой же вирус попадет в организм, эти антитела свяжутся с поверхностным вирусным белком и нейтрализуют вирус. *в* — Синтетический пептид, имитирующий часть поверхностного белка вирусной частицы, также вызывает образование антител, способных нейтрализовать вирус.

ется механизм иммунного ответа, одно из проявлений которого — пролиферация клеток плазмы крови, секретирующих антитела против определенных участков поверхности вируса. Антитела связываются с этими участками и покрывают поверхность вирусной частицы. Они не дают вирусу присоединиться к клеточным рецепторам и тем самым нейтрализуют его.

Сайт, который узнается специфическим лимфоцитом и с которым впоследствии взаимодействует антигенсвязывающий центр специфического антитела, называется антигенной детерминантой. Обычно это небольшой фрагмент поверхностного белка вируса. Если это так, то нельзя ли получить антитела, способные инактивировать вирус, вводя в качестве вакцины те пептиды, которые служат антигенными детерминантами? Представление о том, что даже небольшая часть инфекционного агента может играть роль вакцины, возникло более 45 лет назад: У. Гёбель из Рокфеллеровского института медицинских исследований продемонстрировал, что иммунизация мышей полисахаридами клеточной стенки бактерий, вызывающих пневмонию, защищает животных от последующей инфекции. В 60-е годы Р. Меррифилд (Рокфеллеровский университет) разработал автоматизированный метод соединения аминокислот в пептиды. Это давало возможность получать иммуногенные пептиды заданной структуры.

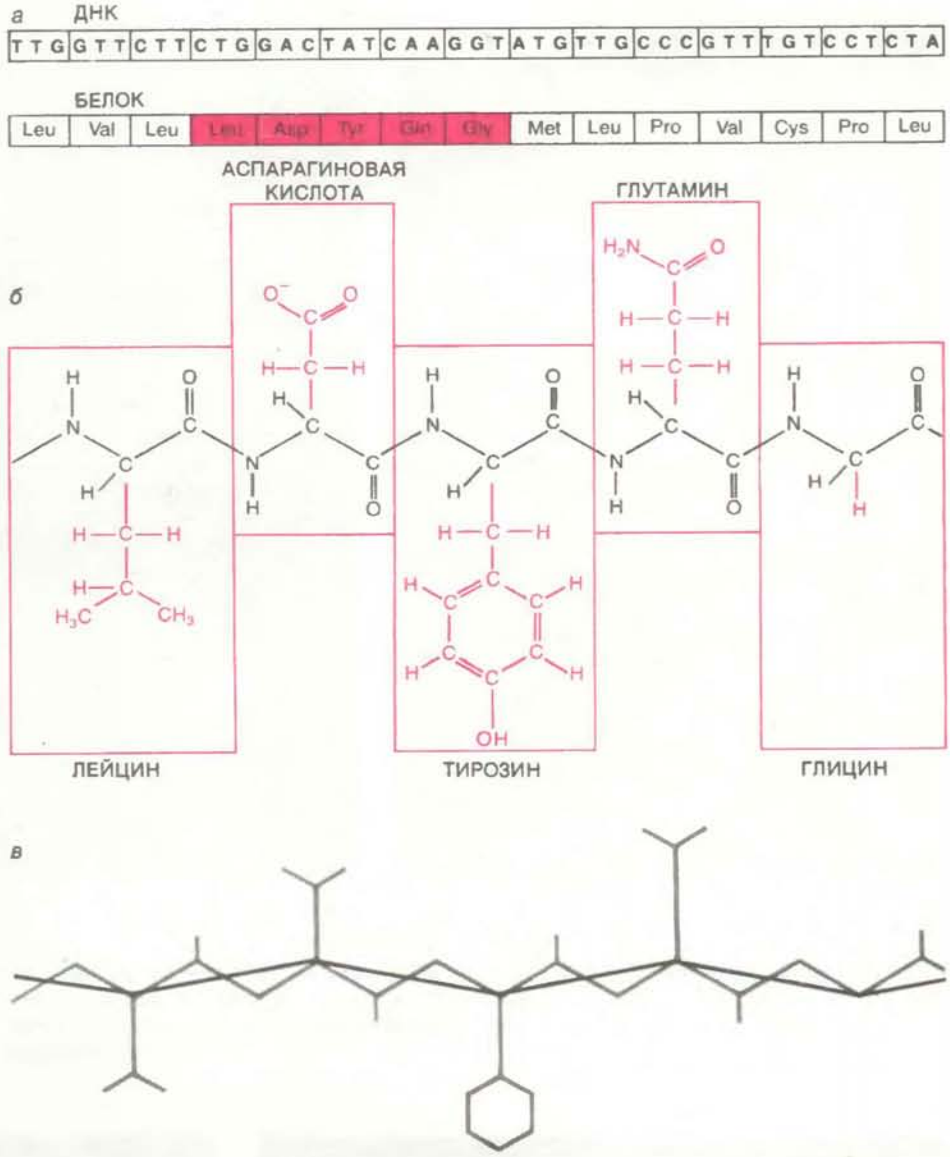
Порой далеко не просто определить аминокислотную последовательность конкретных пептидов в составе длинной белковой молекулы, однако исследователи нашли способ избегать таких затруднений. Аминокислотная последовательность белка задается последовательностью нуклеотидов в нуклеиновой кислоте (ДНК или РНК), кодирующей этот белок. Несколько лет назад были разработаны замечательные новые методы определения нуклеотидной последовательности ДНК. Стало возможным очень быстро определять полную последовательность ДНК вируса (или ДНК-копии вирусной РНК). Поэтому в принципе можно идентифицировать последовательность, которая кодирует поверхностный белок вируса, и транслировать ее согласно генетическому коду в аминокислотную последовательность белка. Затем по методу Меррифилда можно синтезировать короткие пептиды из разных участков этой белковой молекулы, инъецировать их лабораторным животным и выяснить, будут ли они вызывать образование антител, связывающихся с естественным вирусным белком и нейтрализующих вирус.

Однако, прежде чем осмысленно разрабатывать конкретные вакцины, предстояло разобраться в природе ан-

тигенной детерминанты, понять, можно ли имитировать ее с помощью короткого синтетического пептида. Связывание антител с участками белковых молекул определяется их электрическим зарядом и формой. Остов белковой молекулы — это цепочка из атомов углерода и азота, от нее торчат в разные стороны боковые группы аминокислот, каждой из которых свойственны лишь ей присущие форма и распределение зарядов. Последовательность аминокислот (т.е. первичная структура белка) и их близкие взаимодействия определяют, какова вторичная структура данного участка полипептидной цепи: будет ли это спираль или излом. Взаимодействия между отдельными участками белка на уровне вторичной структуры в свою очередь определяют, какова третичная структура белка. Белки, как правило, имеют сложную третичную структуру.

Следовательно, возможны два вида антигенных детерминант. Если антигенсвязывающий центр молекулы антитела распознает боковые группы аминокислот, расположенных в полипептидной цепи друг за другом, говорят о непрерывной антигенной детерминанте. Составная же антигенная детерминанта складывается из боковых групп аминокислот, удаленных друг от друга вдоль полипептидной цепи, но сближенных в пространстве благодаря третичной структуре.

Еще два-три года назад те знания и представления, которыми располагала иммунология, не позволяли надеяться, что синтетические пептиды смогут имитировать естественные антигенные детерминанты интактных вирусных белков. Во-первых, оказалось, что иммунный ответ на вирусный белок при естественной инфекции направлен на небольшое число антигенных детерминант, которые представляют собой не всю белковую молекулу, а лишь отдельные ее участки. В частности, есть данные о том, что для иммуногенности белка имеет значение лишь один небольшой участок на каждые 50 аминокислот. Во-вторых, большинство антигенных детерминант, по-видимому, являются составными. Чтобы получить пептид, заменяющий составную детерминанту, нужно синтезировать очень длинный отрезок белковой молекулы; технически это возможно, но неэффективно. Если бы и удалось синтезировать длинный пептид, маловероятно, что в отсутствие остальной части молекулы белка он свернется в «правильную» третичную структуру. Считалось, что боковые цепи отдельных аминокислот короткого пептида не могут самопроизвольно сориентироваться так, чтобы имитировать структуру даже относительно более простой непрерывной детерминанты.



НУКЛЕОТИДНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ участка вирусной ДНК кодирует последовательность аминокислот белка вируса (*а*). Если известна последовательность ДНК, можно предсказать аминокислотную последовательность и выбрать из нее те короткие отрезки (пептиды), которые должны вызывать образование антител. На рисунке *б* изображена химическая структура части пептида с рисунка *а*. Остов полипептидной цепи построен из атомов углерода (C), азота (N), водорода (H) и кислорода (O). Боковые цепи аминокислот (*цветные*) отходят от так называемых α-углеродных атомов. На полученных с помощью ЭВМ изображениях, которые мы приводим в этой статье, молекула показана либо упрощенно, как остов из α-углеродных атомов (*черный цвет*), либо с деталями на атомном уровне, но без атомов водорода (*серый цвет*).

ДВА ГОДА назад мы бросили вызов этой пессимистической точке зрения. Мы решили синтезировать в виде коротких фрагментов почти всю молекулу вирусного белка и выяснить, не реагируют ли с целым белком антитела против этих пептидов. Для этой цели мы избрали молекулу гемагглютинина одного из штаммов вируса гриппа. Это гликопротеид (т.е. белок, к которому присоединена углеводная цепь), он образует шипы на поверхности сферической оболочки вируса. Гемагглютининовые шипы выполняют важные функции. Они участвуют в связывании вируса с клеткой, а также вызывают слипание красных кровяных клеток (что и отражено в их названии). Именно гемагглютининовые шипы существенны для индукции образования антител против вируса гриппа.

Гемагглютинин вируса гриппа очень подходил для наших целей еще и потому, что У. Фирс и его коллеги (Гентский университет, Бельгия) определили нуклеотидную последовательность гена, кодирующего гемагглютинин, и по ней предсказали его аминокислотную последовательность, а А. Уилсон и Дж. Скехель (Национальный институт медицинских исследований, Лондон) с помощью рентгеноструктурного анализа определили третичную структуру белка. Благодаря этому мы

могли каждый синтезированный нами пептид соотносить с той или иной областью вторичной и третичной структуры. Молекула гемагглютинина состоит из двух полипептидных цепей, которые обозначают *HA*1 и *HA*2; каждый гемагглютининовый шип состоит из трех таких молекул. Главной мишенью иммунной системы, по-видимому, является цепь *HA*1, поэтому мы начали с нее.

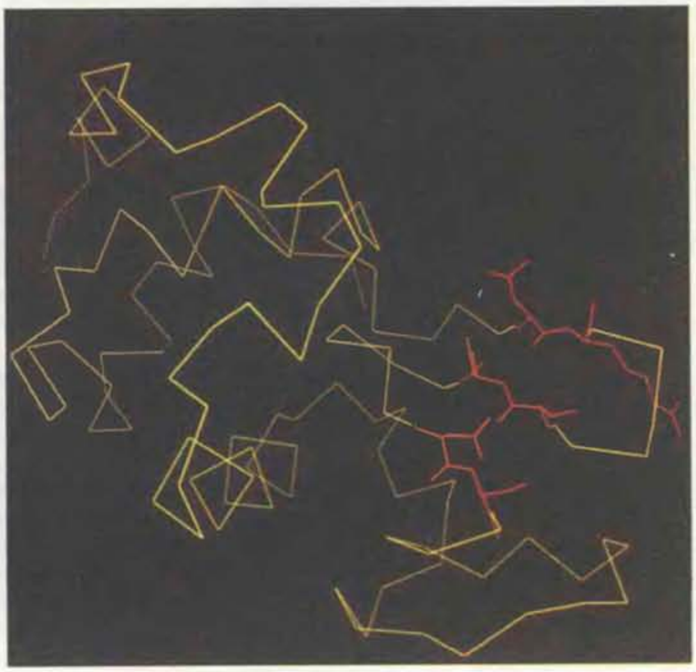
Мы синтезировали 20 пептидов, суммарная длина которых составляла около 75% длины цепи *HA*1; они дали нам целый набор вторичных и третичных структур. Затем каждый пептид присоединили к крупному белку-носителю и инъецировали кроликам. Сыворотку от каждого кролика исследовали на присутствие в ней антител, способных связываться с данным пептидом, с гемагглютинином и с цельной вирусной частицей. Против всех 20 пептидов образовывались антитела; 15 видов таких антител связывалось также с гемагглютинином и 17 видов — с вирусом гриппа. Другими словами, пептид почти из любого участка молекулы вирусного белка может вызывать образование антител, которые узнают целый белок как в очищенном виде, так и на поверхности вирусной частицы. Притом иммуногенность нативного белка меньше потенциальной иммуногенности его частей. По-видимому, требуется только, чтобы синтезированный пептид соответствовал участку, располагающемуся не в глубине свернутой молекулы белка, а на ее поверхности, где антитела могут с ним контактировать.

Есть несколько приемов, позволяющих определить, находится ли данный пептид на поверхности белковой глобулы — более строго, доступен ли пептид (или его часть) для молекул антител в крови животного. Один из наиболее точных методов заключается в том, что с помощью ЭВМ строят графическое изображение участков молекулы белка, доступных для молекул растворителя (см. рисунок на с. 5). Сначала получают изображение исследуемого пептида (с боковыми группами всех его аминокислот), причем именно в той конформации, которую он принимает в целом белке. Затем моделируют с помощью ЭВМ, как по поверхности пептида будет катиться сфера размером с молекулу воды. Сфера проникает в углубления на поверхности молекулы везде, где ей это позволяет соотношение вандерваальсовых радиусов молекулы воды и атомов боковых групп (вандерваальсов радиус примерно соответствует расстоянию, на котором еще действует электрический заряд). В каждом месте контакта сферы с пептидом ставится точка; таким образом получается карта участков пептида, доступных молекулам растворителя.

Этот подход эффективен, но требует детального, с атомным разрешением знания структуры вируса и его белков, а такие данные есть лишь для немногих вирусов. К счастью, зная только аминокислотную последовательность белка, можно предсказать, какие участки будут располагаться на поверхности молекулы. Большинство аминокислот либо гидрофильны (т.е. притягивают воду), либо гидрофобны (т.е. отталкивают воду). Гидрофобные аминокислоты обычно сокрыты в сердцевине белковой глобулы, а гидрофильные — расположены на ее поверхности. Наиболее гидрофильны четыре аминокислоты — лизин, аргинин, аспарагиновая кислота и глутаминовая кислота. Считается, что те участки молекулы белка, в которых этих аминокислот относительно много, предпочитают водное окружение и поэтому располагаются на поверхности.

Как я уже отмечал, в наших экспериментах получалось, что практически любой пептид с поверхности белка вируса гриппа вызывает образование антител, способных реагировать с целым белком. Но разве может изолированный пептид принять ту же конформацию, что и в целой, сложным образом свернутой молекуле белка? Не исключено, что правильная форма молекулы складывается именно благодаря отсутствию пространственных ограничений. Пептид небольшого размера, введенный в кровоток животного, находится в растворе, где ничто не мешает ему все время изменять свою форму, изгибаясь и закручиваясь. Сталкиваясь при этом с лимфоцитами, он вы-

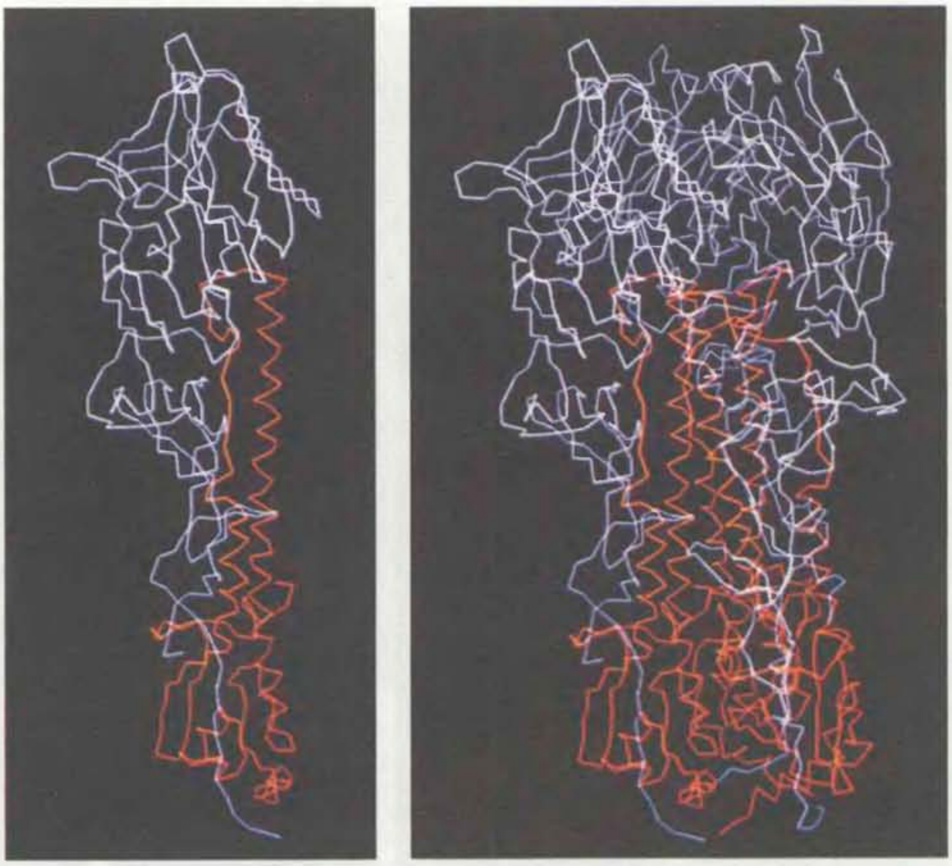




АНТИГЕННЫЕ ДЕТЕРМИНАНТЫ — это те участки на поверхности белковой молекулы, которые узнают антитела и с которыми они связываются. На этом рисунке мы приводим полученное с помощью ЭВМ изображение части полипептидной цепи фермента лизоцима (*желтая*); показаны два типа антигенных детерминант с боковыми группами аминокислот (*оранжевые*). Непрерывная детерминанта (*слева*) — это группа соседних в полипептидной цепи аминокислотных остатков. Составная детерминанта (*справа*) формируется из нескольких участков, удаленных друг от друга вдоль полипептидной цепи, но сближенных в пространстве благодаря третичной структуре белка.

зывает образование антител против каждой конформации в отдельности. Одна из конформаций свободного пептида точно совпадает с конформацией соответствующего участка в молекуле белка. (Интересно отметить, что во многих случаях короткие пептиды более эффективны в смысле образования антител против вирусного белка, чем сам этот белок в очищенном виде. Возможно, после выделения и очистки белок приобретает конформацию, отличную от той, которую он имел на поверхности вирусной частицы, где его окружали такие же молекулы.)

ЕСЛИ ПРАКТИЧЕСКИ любой участок поверхности молекулы вирусного белка может вызывать образование антител против этого белка, значит, можно конструировать вакцины, а не получать их в виде массы разнообразных антител, которые дает иммунная система организма в ответ на введение целых вирусных частиц. Этой задачей занимается ряд лабораторий. Первая вакцина, которую мы разработали, была против ящура. Ящур — давний бич животных, в особенности крупного рогатого скота. Это заболевание до сих пор остается серьезной проблемой скотоводства практически повсюду, за исключением Австралии и Северной Америки; ежегодно тратится несколько миллиардов доз вакцины с убитым вирусом. Оболочка вируса ящура построена из 240 молекул белков четырех типов — *VP*1, *VP*2, *VP*3 и *VP*4, по 60 молекул каждого типа. Пытались получить вакцину на основе белка *VP*1, который считается главной мишенью антител. Однако оказалось, что у целого белка иммуногенность низкая. Поэтому в сотрудничестве с Ф. Брауном и Д. Роуландсом из Института по изучению вирусов животных (Великобритания) мы начали эксперименты с целью синтезировать фрагмент белка *VP*1 более иммуногенный, чем целый белок.



МОЛЕКУЛА ГЕМАГГЛЮТИНИНА вируса гриппа состоит из двух полипептидных цепей: *HA* 1 (*фиолетовая*) и *HA* 2 (*оранжевая*). Трехмерная модель молекулы (*слева*) построена по данным рентгеноструктурного анализа, полученным А. Уилсоном и Д. Уили (Гарвардский университет) и Дж. Скехелем (Национальный институт медицинских исследований, Лондон); показан лишь остов белка. Тример гемагглютинина (*справа*) содержит три цепи *HA* 1 и три цепи *HA* 2. Тримеры образуют шипы на сферической оболочке вирусной частицы.

Последовательность 213 аминокислот белка *VP*1 уже была известна из нуклеотидной последовательности, но детальная, с атомным разрешением, структура белка до сих пор не выяснена. Однако оказалось достаточно проанализировать аминокислотную последовательность, чтобы обратить внимание на участок между 141-й аминокислотой и 160-й. В этом участке очень много гидрофильных боковых групп аргинина, лизина и аспарагиновой кислоты, следовательно, он скорее всего располагается на поверхности молекулы белка. В местах изломов полипептидной цепи (излом, резкий изгиб может вносить вклад в иммуноспецифичность молекулы) часто находится пролин, а между 141-й и 160-й аминокислотами как раз есть два остатка пролина.

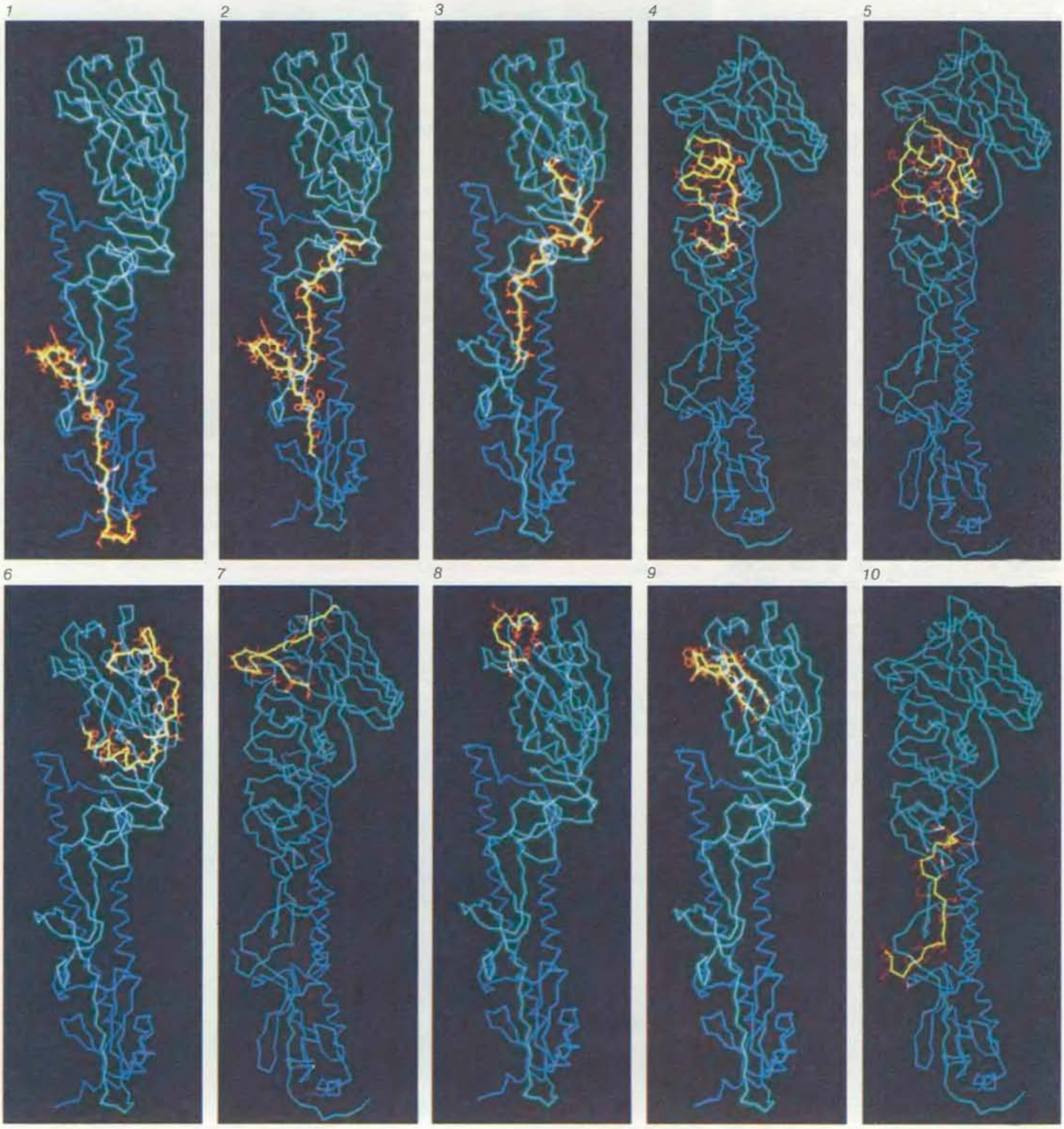
Важно отметить, что в области между 141-й и 160-й аминокислотами последовательности белка *VP*1 у разных штаммов вируса заметно различаются. Вирусу нужно уметь ускользать от действия иммунной системы хозяина достаточно долго, чтобы успеть заразить его клетки и дать потомство. Для этого он постоянно изменяет антигенные детерминанты на своей оболочке, так что антитела, образовавшиеся ранее против исходного штамма, не могут связаться с измененной формой вируса и нейтрализовать ее. Говоря менее телеологическим языком, естественный отбор позволяет выжить тому штамму вируса, в котором благодаря мутации произошло иммунологически значимое изменение. Стало быть, последовательность для синтетической вакцины следует искать именно в вариабельном участке молекулы, мутации в котором меняют его иммуноспецифичность.

Мы синтезировали несколько пептидов белка *VP*1 вируса ящура, в том числе и пептид с аминокислотами со 141-й по 160-ю. Каждый пептид присоединили к молекуле белка-носителя и ввели кроликам. Мы убедились, что пептиды индуцируют образование антител против самих себя, и поставили биопробу, чтобы посмотреть, нейтрализуют ли эти антитела вирус. Когда мышам вместе с вирусом вводили антисыворотку кролика против пептида 141—160, она эффективно инактивировала вирус — с антисывороткой мыши получали так называемый пассивный иммунитет. Существуют три основных штамма вируса ящура; мы работали с вирусом типа *O*. Антисыворотка на штамм *O* нейтрализовала два других штамма гораздо хуже. Это подтверждает представление о том, что изменения в определенном участке белка *VP*1 позволяют вирусу избегать действия антител против другого штамма.

Главное требование ко всякой вакцине заключается в том, что она должна создавать не только пассивный, но и активный иммунитет, то есть индуцировать в организме образование собственных антител на случай возможного заражения таким же вирусом в будущем. Мы иммунизировали морских свинок небольшими дозами (200 мкг) синтетической вакцины и через 35 дней ввели мощную дозу вируса, в 10 000 раз бо́льшую, чем достаточно для заражения половины животных в

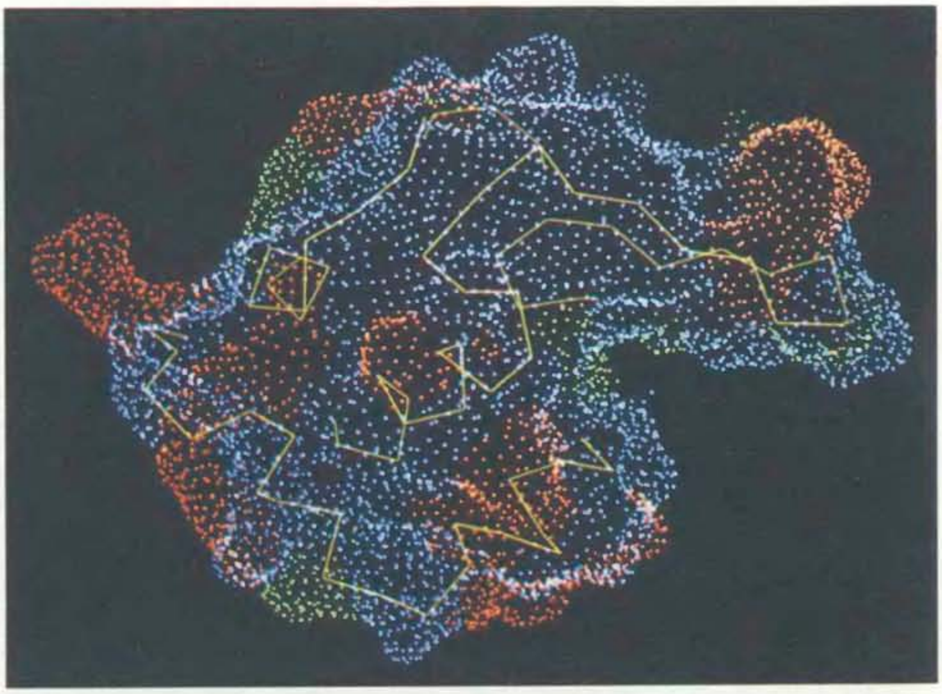
опытной группе. Ни одна из иммунизированных морских свинок не заболела, а неиммунизированные заболели все. Сейчас в Великобритании аналогичным путем исследуют активный иммунитет у крупного рогатого скота. Мы уже установили, что у скота, иммунизированного пептидом, вырабатываются антитела в концентрации, достаточной для защиты от вирусного заражения. Для практического применения нужна, конечно, вакцина, которая защищает от всех трех штаммов вируса. Мы синтезировали пептиды, соответствующие одному и тому же участку (141—160) белка *VP*1 вирусов типа А и вирусов типа С (аминокислотные последовательности вирусов типов А и С несколько различаются). С этими пептидами были проделаны биопробы на мышах; опыты показали, что они действительно защищают от вирусов типа А и типа С. По-видимому, не за горами создание вакцины сразу против трех штаммов вируса ящура.

Итак, в принципе мы можем осу-



ПЕПТИДЫ, общая длина которых составляет 75% длины молекулы *НА* 1, исследовали на способность вызывать образование антител против самого пептида, против молекул гемагглютинина и против целых частиц вируса гриппа. На рисунке изображены 10 из 20 изученных пептидов (*желтые*), их боковые группы (*оранжевые*), остальная часть молекулы *НА* 1 (*светло-синяя*), молекула *НА* 2 (*темно-синяя*); в каждом случае мономер гемагглютинина повернут так, чтобы пептиды было лучше видно. Большинство этих пептидов вызывало у кроликов образование антител, связывающихся с вирусными частицами.

ществить все методические этапы от последовательности нуклеиновой кислоты до синтетической вакцины, теперь настала пора изучить отдельные аспекты разработки вакцин более детально. В настоящее время ищут новые, более тонкие способы определения тех участков вирусных белков, которые наиболее чувствительны к действию антител; кроме того, необходимо добиться увеличения длительности иммунитета.



КАРТА ПОВЕРХНОСТИ МОЛЕКУЛЫ ЛИЗОЦИМА, доступной растворителю, получена с помощью метода, показанного на иллюстрации к с. 5. Цвет точки отражает природу аминокислоты: осно́вные аминокислоты лизин и аргинин оранжевые, кислые аминокислоты аспарагиновая и глутаминовая зеленые, все остальные синие. Осно́вные и кислые аминокислоты наиболее гидрофильны (т.е. притягивают воду). Видно, что они склонны располагаться на поверхности. Если детальная структура белка неизвестна, можно просто принять, что пептиды, богатые гидрофильными аминокислотами, должны быть иммуногенными.

В МОЛЕКУЛАХ вирусных белков имеются участки двух видов. В одних возможны вариации — это особенно важно, когда вирусу нужно ввести в заблуждение иммунную систему хозяина. Вакцина, которая индуцирует образование антител против такого участка (например, против полипептидной цепи *HA*1 вируса гриппа или против белка *VP*1 вируса ящура), обычно иммуногенна, но для того, чтобы она была эффективна против каждого нового штамма, ее придется «подгонять». Участки другого типа выполняют какую-либо жизненно важную функцию, присущую всем штаммам вируса, и их структура неизменна. Аминокислотная последовательность таких участков в высшей степени консервативна, т.е. у всех штаммов данного вируса она практически одна и та же. Если такой участок доступен действию антител, вакцина, которая их индуцирует, будет универсальной — от нее не сможет уберечься ни один штамм вируса.

Вирус гриппа недаром пользуется дурной славой: он ловко ускользает от действия иммунной системы, изменяя поверхностные участки гемагглютинина. Изменения затрагивают цепь *HA*1 (в опыте, о котором я рассказывал выше, мы исследовали иммуногенность пептидов именно из *HA*1). Гемагглютинин, однако, выполняет еще по меньшей мере две неизменные функции. Во-первых, на поверхности его молекулы имеется впадина, в которой связывается сиаловая кислота — компонент клеточной стенки; с этого и начинается инфекционный процесс. Во-вторых, гемагглютинин вызывает слияние мембран зараженных клеток, что увеличивает инфекционность вируса. П. Чоппин и его коллеги из Рокфеллеровского университета и М. Уотерфилд из Государственного фонда по изучению рака получили данные, позволяющие предполагать, что консервативный участок в начале цепи *HA*2 ответствен за слияние мембран или по меньшей мере участвует в нем.

Обычно иммунная система, так сказать, не замечает этого участка; при нормальной инфекции антител против него не образуется. Если бы консервативные участки были в норме иммуногенными, против них образовывались бы антитела. Тогда, один раз переболев гриппом, человек или животное приобретали бы иммунитет против всех вариантов вируса гриппа, и это заболевание просто исчезло бы. Мы решили выяснить, нельзя ли получить антитела сразу против нескольких штаммов вируса, используя для иммунизации пептид из консервативного участка.

С. Александер, А. Александер и Н. Грин синтезировали несколько пептидов из начального участка полипептидной цепи *HA*2 вируса гриппа типа А подтипа H_3. При тестировании в культуре клеток антисыворотки кроликов, полученные против этих пептидов, эффективно инактивировали вирус А, H_3. Следовательно, даже если начальный участок цепи *HA*2 в норме и не является иммуногенным, он доступен специфическим антителам. Более того, эти же антисыворотки инактивировали вирусы других подтипов и даже другого типа (а именно вирус гриппа В). Антитела против целой вирусной частицы подтипа H_3, напротив, не инактивировали вирусы других подтипов. И конечно, животные, иммунизированные целым вирусом гриппа какого-либо одного подтипа, не становились устойчивыми к заражению вирусами других подтипов. Прежде чем говорить о практическом применении анти-*HA*2-антител, нужно исследовать их более подробно, но первые результаты уже многое обещают и позволяют возлагать большие надежды на вакцины, в которых используются консервативные участки молекул.

Итак, один из способов создания вакцины заключается в том, чтобы выбрать тот фрагмент белковой молекулы, который, по теоретическим соображениям, может играть роль антигенной детерминанты и синтезировать его. Есть и другой путь. Можно попытаться имитировать антигенные детерминанты, которые уже известны из эпидемиологии серологических форм. В серологии вирусы классифицируются по свойствам антител, которые они индуцируют. Серологические исследования позволяют выявить различные типы и подтипы вирусов; о некоторых из них я упоминал, рассказывая о вирусах ящура и гриппа.

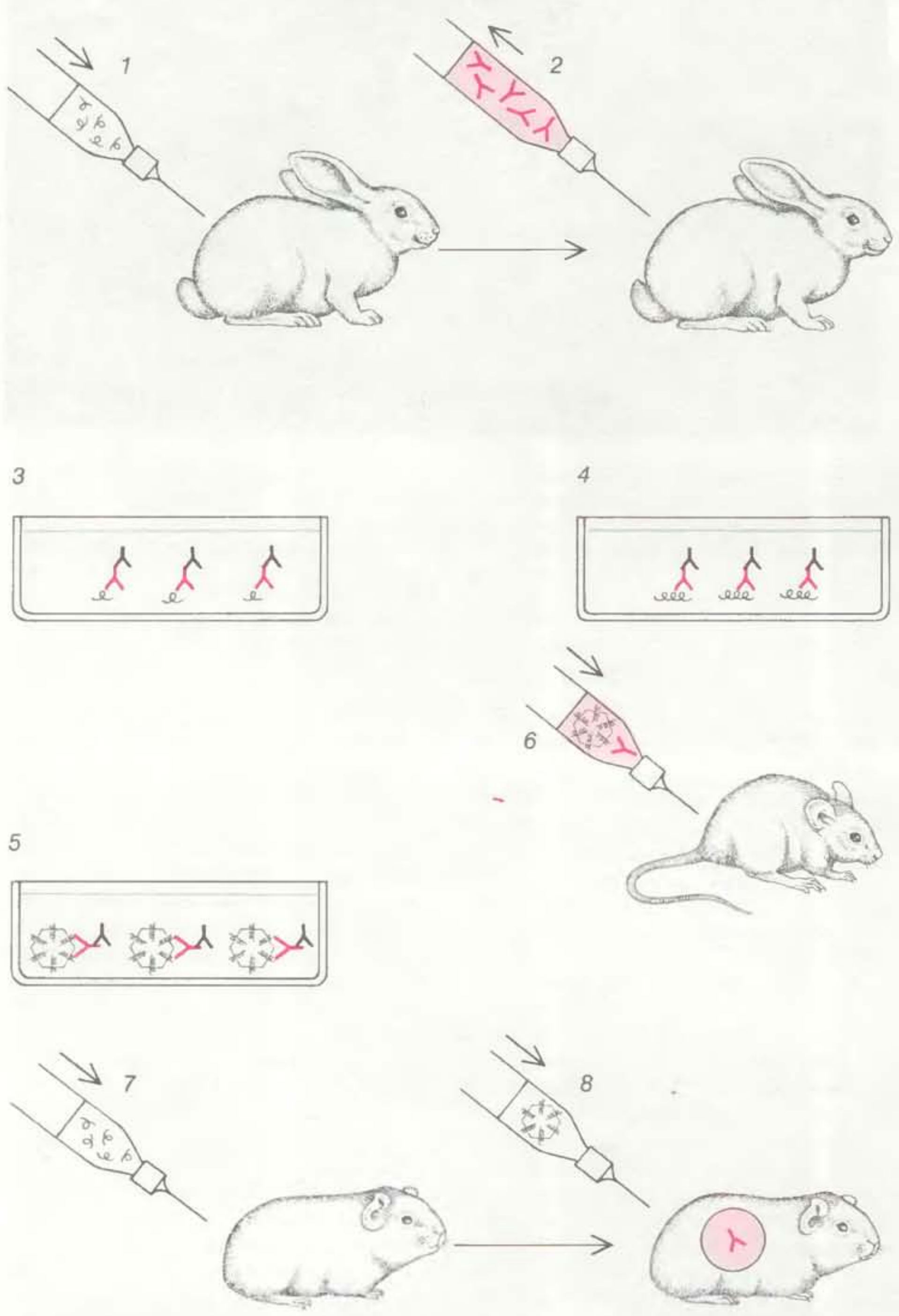
Чтобы определить тип вируса, нужны длительные исследования. Вирус, вызывающий данное заболевание, выделяют из организмов больных людей или животных в разных районах и в разное время. Антитела из крови больных испытывают против набора вирусов; разные антитела реагируют с разными штаммами вируса, поскольку они «узнают» разные антигенные детерминанты. На основании особенностей взаимодействия антител с различными вирусами из этого набора иммунологи могут серологически описать данный тип вируса с точки зрения ан-

тигенных детерминант и определить, какие детерминанты «работали» у данного больного или во время конкретной эпидемии. Однако до недавнего времени детерминанты оставались абстрактными маркерами, их обозначали буквами или цифрами, а о молекулярной природе ничего не знали. Поэтому трудно было предположить, что данные по эпидемиологии серологических типов окажутся полезными при конструировании вакцин. Конечно, возможна ситуация, когда синтетический пептид не может имитировать структуру, отвечающую данному серологическому маркеру: например, если эта структура составная. Мы решили попытаться получить вакцину против одного из заболеваний печени — гепатита *B*, — синтезировав пептиды, эквивалентные уже известным серологическим детерминантам, характерным для разных типов вируса.

В СЛУЧАЕ гепатита *B* главной мишенью антител, т.е. поверхностным антигеном вируса, служит, по-видимому, белок оболочки вирусной частицы. Серологические исследования показывают, что у разных штаммов вируса гепатита *B* есть одна общая детерминанта, ее обозначают *a*. Каждый штамм имеет еще две детерминанты: *d* (или альтернативную ей *y*) и *w* (или альтернативную ей *r*). Это значит, что возможны четыре типа вируса: *adw*, *ayw*, *adr* и *ayr*. В сотрудничестве с Дж. Герином (Медицинская школа Джорджтаунского университета) и Р. Пёрселлом (Национальный институт аллергии и инфекционных заболеваний) мы задались целью синтезировать пептиды, которые бы индуцировали образование антител против детерминант *d* и *y*. Молекула поверхностного антигена вируса гепатита *B* состоит из 226 аминокислот, т.е. довольно велика, поэтому нужны были какие-то указания, откуда начать поиск. Аминокислотная последовательность антигена была проанализирована для трех штаммов вируса. Оказалось, что существенные различия есть только в одном месте — между 110-й аминокислотой и 140-й. Различие между детерминантами всегда обусловлено различиями в последовательностях аминокислот, стало быть, нужные нам пептиды следовало искать именно в этом участке.

Мы синтезировали несколько пептидов, которые вместе соответствовали участку 110—140, и получили антитела против каждого из них. Один пептид из 13 аминокислот (со 125-й по 137-ю) индуцировал образование антител, которые связывались только с вирусами, несущими детерминанту *y*. Это означало, что детерминанта *y* и альтернативная ей детерминанта *d* должны входить в пептид 125—137. Синтезировав *y*, мы получали возможность синтезировать и *d* по уже известной нам последовательности участка 125—137 у другого вируса. Мы выбрали последовательность, которая отличалась от предыдущей всего тремя аминокислотами: на месте двух остатков треонина и остатка тирозина в ней стояли соответственно пролин, аспарагин и фенилаланин. Антитела против этого пептида связывались только с вирусами, несущими детерминанту *d*.

Возможно, различие между детерминантами *y* и *d* определяется менее чем тремя заменами аминокислот.



СИНТЕТИЧЕСКИЙ ПЕПТИД вводят кролику (*1*), чтобы выяснить, вызывает ли он образование антител. Сыворотку от кролика (*2*) сначала исследуют на присутствие антител против введенного пептида (*3*), против целого вирусного белка (*4*) и против интактной вирусной частицы (*5*). О связывании антипептидных антител (*цветные*) судят по активности фермента, «пришитого» ко вторым антителам (*черные*), которые получены против антипептидных антител. Чтобы определить способность антипептидных антител нейтрализовать вирус, т.е. создавать пассивный иммунитет, мышам одновременно вводят вирус и антисыворотку (*6*): если мышь не заболевает, это означает, что вирус нейтрализован. Чтобы определить способность синтетического пептида служить вакциной, создающей активный иммунитет, морским свинкам вводят пептид (*7*), а затем вирус (*8*). Если животное выживает (неиммунизированные контрольные животные гибнут), это доказывает эффективность вакцины.

Сколько замен необходимо и достаточно, мы выясним, получив пептиды только с одной из трех замен. Пока что можно сделать два вывода. Во-первых, различия между разными штаммами одного и того же вируса определяются очень небольшими изменениями в аминокислотном составе белка. Во-вторых, синтетические пептиды имитируют особенности, выявляющиеся в серологических исследованиях, следовательно, при разработке синтетических вакцин можно будет воспользоваться данными серологии. Будем надеяться, что это даст возможность справляться с неожиданными эпидемиями, вызванными мутантными штаммами уже известных вирусов.

Чтобы синтетические вакцины могли заменить обычные вакцины, необходимо научиться повышать уровень иммунитета, который они вызывают, и продлевать его. Как я уже упоминал, ослабленные вирусы удобны тем, что сохраняют способность размножаться и, таким образом, все время поставляют антиген для иммунной системы организма, которая в результате вырабатывает все больше и больше антител. Синтетический пептид, как и убитый вирус, не размножается, поэтому к вакцине приходится добавлять так называемые адъюванты — они повышают уровень иммунитета и увеличивают его длительность.

Один из наиболее эффективных адъювантов — адъювант Фрейнда — представляет собой эмульсию из минерального масла и воды, смешанную с убитыми микобактериями (к микобактериям относится, в частности, возбудитель туберкулеза). Механизм его действия неизвестен: предполагают, что адъюванты действуют двояко. Эмульсия, вероятно, служит как бы ловушкой антигена, создавая своего рода депо, из которого антиген медленно высвобождается, что имитирует постоянный синтез антигена реплицирующимся вирусом. Убитые бактерии привлекают к этому депо различные клетки, участвующие в иммунном ответе. Однако скопление таких клеток неизбежно вызывает раздражение в месте инъекции и даже абсцесс. Мощные адъюванты типа адъюванта Фрейнда нельзя вводить человеку — вместо них в вакцины с убитыми вирусами добавляют квасцы или гидроокись алюминия. Антиген адсорбируется на частицах гидроокиси алюминия, а в организме начинает постепенно освобождаться. По-видимому, для некоторых синтетических вакцин квасцы вполне подойдут, но в большинстве случаев потребуется что-то более эффективное. Сейчас многие исследователи работают над созданием эффективных и безопасных адъювантов.

Обнадеживающих результатов достигли Ф. Одибер и Л. Шедид (Институт Пастера) и Р. Арнон и М. Села (Вейсмановский институт). Им удалось получить антитела, обеспечивающие пассивный иммунитет к дифтерийному токсину. Они синтезировали три пептида, соответствующих участкам молекулы токсина. Прежде чем вводить синтетический пептид в организм, нужно соединить его с белком-носителем. (В наших опытах носителем служил гемоцианин — дыхательный фермент моллюска из семейства Fissurellidae; этот белок часто используют для такой цели.) Одибер, Шедид, Арнон и Села предложили комбинацию из белка-носителя и адъюванта. Они присоединили синтезированные пептиды дифтерийного токсина к носителю, а носитель — к препарату мембран микобактерий. Мембраны играли роль адъюванта и сильно увеличивали иммуногенность пептидов.

Я УЖЕ говорил, что антитела против пептидов наверняка найдут применение как инструменты исследования. В первую очередь они пригодятся для выявления и идентификации белков — продуктов изучаемых генов. Мы достигли большой точности в получении рекомбинантных ДНК и умеем определять последовательности нуклеиновых кислот, поэтому часто складывается такая ситуация, когда полная последовательность гена, кодирующего какой-то белок, известна, а сам белок не найден и его нужно обнаружить. По последовательности ДНК и РНК можно предсказать аминокислотную последовательность искомого белка. Если затем выбрать в ней небольшой участок и синтезировать такой же пептид, антитела против этого пептида послужат специфической пробой для обнаружения данного белка. Например, таким способом были идентифицированы белки, кодируемые некоторыми онкогенами (онкогены — это гены некоторых вирусов, индуцирующие образование опухолей). Определили и локализацию онкобелков в клетке.

Отличительная черта антипептидных антител — их заранее известная специфичность. Они представляют собой антитела не против всего белка, а против определенной известной экспериментатору его части. Образовавшиеся в клетках иммунной системы лабораторного животного антитела можно легко очистить, пропустив один раз через колонку с иммуносорбентом. С помощью антител удобно исследовать детали структуры белков, а также взаимосвязь структуры и функций. Например, антитела против какого-то определенного участка молекулы белка, связываясь только с такими участками, будут избирательно подавлять их функцию; имея набор антител против различных участков исследуемого белка, можно определить его активный центр (или центры) и соотнести функции белка с конкретными структурами. Многие пептидные гормоны синтезируются первоначально в виде длинных полипептидных цепей, которые затем расщепляются специальными ферментами на более мелкие молекулы, каждая из которых представляет собой отдельный гормон со специфической активностью. Антипептидные антитела позволят разобраться в такой смеси и поставить в соответствие отдельные гормональные активности с конкретными пептидами. Наконец, используя антитела с заранее известной специфичностью, можно выяснить, какая часть гена когда экспрессируется. В генах млекопитающих кодирующие участки (экзоны) прерываются некодирующими, так называемыми вкрапленными чередующимися последовательностями (интронами). В генах вариабельных белков в разное время экспрессируются разные экзоны. Например, при иммунном ответе последовательно образуются разные формы антител. Теперь мы можем синтезировать пептиды, соответствующие отдельным экзонам, получить против них антитела и определить, какой экзон функционирует на каждой стадии иммунного ответа.

Кремниевые мастерские

Специалист фирмы Intel Corporation Г. Мур отмечает, что для развития микроэлектроники характерны две закономерности — ежегодное увеличение в два раза количества схемных элементов в одном кремниевом чипе и пропорциональная зависимость стоимости разработки микросхем от их сложности. Например, на разработку большой интегральной схемы, выполняющей функции микропроцессора и содержащей более 450 тыс. элементов, требуется не менее двух лет. Стоимость конструирования такого устройства может превышать производственные затраты на изготовление всей партии чипов данной модели.

Рост стоимости конструкторской разработки интегральных схем привел к появлению новых, более производительных способов их создания. Для этих целей, кроме того, широко стали внедряться системы автоматизированного проектирования на базе ЭВМ (САПР), которые избавили разработчика от выполнения малопроизводительных операций. Эти нововведения могут оказать существенное влияние на промышленное производство микроэлектронной продукции. Разработчики систем, которые в настоящее время создают электронное оборудование из готовых чипов, поставляемых фирмами, специализирующимися на выпуске полупроводниковой продукции, теперь, возможно, сами станут конструировать нужные им интегральные схемы.

В настоящее время специальное назначение электронных систем в основном определяется программными средствами. Обычное электронное устройство состоит из нескольких программируемых чипов большой степени интеграции, таких, как микропроцессоры и запоминающие устройства, и нескольких более простых чипов, необходимых для обслуживания сложных устройств; все эти узлы собираются вместе и монтируются на печатной плате. Разработчик системы приспосабливает ее для определенных задач путем подбора входящих в систему компонентов и схемы их взаимного соединения, но функции системы определяются главным образом программированием микропроцессоров и блоков памяти. Единственное, чего не в силах сделать разработчик, — это изменить внутреннюю структуру микросхемы.

В последние пять лет появился ряд новых подходов в технологии микросхем, называемых «полузаказными», которые дают возможность системным разработчикам «дорабатывать» некоторые типы интегральных схем и получать нужные характеристики. Одним из таких микроэлектронных устройств, изготовленных по такой технологии, является «вентильная матрица» — чип, на котором несколько тысяч логических элементов (комбинаций схемных элементов, выполняющих простейшие логические функции) соединены рядами в строго определенном порядке. Сами матрицы во всех чипах одни и те же, но конфигурация контактных металлических слоев, которая определяет схему соединения матриц, у разных чипов разная. Разработчик системы выбирает чип с нужной ему схемой межсоединений.

Около 40 компаний уже производят такие вентильные матрицы. Некоторые из них выпускают САПР, которые помогают разработать схему соединений, обеспечивающих наиболее эффективное использование функциональных возможностей каждого чипа. Типичная САПР имеет библиотеку данных для соединения элементарных электронных логических ключей. Разработчик вводит в САПР описание логических функций проектируемого устройства, и система переводит это описание в топологию межсоединений. Топология служит основой для изготовления фотошаблона, который позволяет получить металлические слои.

Итак, вентильная матрица состоит из логических элементов. Составными блоками других микросхем полузаказной технологии, называемых стандартными ячейками, являются логические элементы, запоминающие устройства и центральные процессоры. Восемь фирм, специализирующиеся на производстве полупроводниковой продукции, уже поставляют в продажу кристаллы с набором стандартных ячеек. Изготовители указывают функции каждой ячейки и ее технические характеристики, такие, как, например, время прохождения сигнала. С помощью САПР разработчик может комбинировать различные стандартные ячейки на одном чипе. Он разбивает систему на блоки, функционально соответствующие ячейкам, и составляет так называемые сетевые спецификации, в которых перечислены ячейки и указано, как они должны соединяться. САПР сама размещает эти ячейки и рассчитывает оптимальную схему межсоединений. В конечном итоге система выдает данные, по которым затем формируются все слои чипа.

Стандартные ячейки позволяют разработчику при конструировании микросхем пользоваться большей свободой, чем вентильные матрицы, но и в этом случае он все же лишен возможности вмешаться во внутреннюю структуру самих ячеек. К. Мид из Калифорнийского технологического института и Л. Конуэй из Научно-исследовательского центра Xerox в Пало-Альто разработали конструктивный подход к созданию специализированных интегральных схем, который полностью устраняет зависимость системного разработчика от фирм-изготовителей. На его основе САПР смогут рассчитывать всю схемную компоновку каждого чипа, а не выбирать из банка данных готовые схемные блоки.

Такие автоматизированные системы Мид называет кремниевыми компиляторами по аналогии с компилятором-программой, которая использует машинный язык высокого уровня. Кремниевый компилятор принимает определение функции конструируемого устройства, выраженное на языке высокого уровня, и автоматически преобразует его в топологию чипа в стандартном формате почти так же, как компилятор машинного языка транслирует программы, представленные в терминах процедурно-ориентированного языка, в команды, выполняемые компьютером. Несколько университетов уже разработали такие кремниевые компиляторы, и вскоре появятся первые промышленные образцы.

Кремниевый компилятор позволяет многим инженерам создавать интегральные схемы с параметрами, отвечающими их индивидуальным требованиям. В результате, как отмечает Мид, получат развитие промышленные предприятия нового типа — кремниевые мастерские. Они будут выпускать только чипы, а схемную часть будут разрабатывать сами потребители, подобно тому как это делается с литейными заготовками из металлов, обработку которых осуществляют потребители.

Специализируясь на производстве разнообразных чипов малых объемов, кремниевые мастерские смогут конкурировать с изготовителями микросхем массового производства, выпускающими готовые схемы на чипах относительно больших объемов и ограниченной номенклатуры. Некоторые фирмы — изготовители полупроводниковой продукции уже создают кремниевые мастерские. Они будут совмещать изготовление чипов, ориентированных на последующую разработку схем потребителями, с производством микросхем собственной конструкции. Поскольку кремниевые компиляторы и кремниевые мастерские позволяют получать более дешевые микросхемы с заданными параметрами, электронные системы для различных целей и областей применения можно будет создавать на одном чипе, а не собирать их из печатных плат с большим числом чипов.

Оптический компьютер

Один триллион (10^{12}) операций в секунду — таково быстродействие оптической вычислительной машины, в которой вместо электрических токов будут использоваться световые лучи. Уже создан основной элемент такой машины — оптический аналог транзистора

АЙТЕН ЭЙБРЭХЭМ, КОЛИН Т. СИТОН И С. ДЕЗМОН СМИТ

ЗА ПОСЛЕДНИЕ 40 лет развитие цифровых вычислительных машин было настолько тесно связано с прогрессом электроники, что их, по существу, стали считать электронными устройствами. В действительности же вычислительная машина может выполнять логические и арифметические операции самыми различными способами. В середине 70-х годов появилась идея о потенциальной возможности создания вычислительного устройства, в котором электрические токи можно заменить световыми лучами лазеров. Реализация такой идеи весьма заманчива: быстродействие оптической вычислительной машины почти в 1000 раз будет выше по сравнению с электронным компьютером.

Основными элементами любой цифровой вычислительной машины являются переключающие устройства, которые могут принимать одно из двух устойчивых состояний. Предельная скорость вычислительных операций в машине, помимо прочих факторов, определяется временем перехода переключающего устройства из одного состояния в другое. В электронных вычислительных машинах роль переключающих устройств выполняют транзисторы, но даже самый быстродействующий современный транзистор не может изменить свое состояние менее чем за одну наносекунду (миллиардную долю секунды). Время переключения оптического устройства, аналогичного транзистору, составляет всего одну пикосекунду (тысячную часть миллиардной доли секунды).

Авторы статьи создали экспериментальный вариант оптического транзистора, переключающегося при небольшом изменении интенсивности падающего на него лазерного луча. В основе оптического транзистора, который мы называем трансфазором, лежит свойство некоторых кристаллов изменять показатель преломления при увеличении интенсивности падающего на них света. Как известно, показатель преломления вещества показывает, насколько замедляется свет при прохождении через это вещество. Можно так подобрать свойства кристаллического вещества и длину волны лазерного излучения, что даже при незначительном изменении интенсивности светового луча лазера на входе она будет существенно изменяться на выходе. Созданные нами экспериментальные устройства имеют время переключения, приближающееся к нескольким пикосекундам.

Оптические транзисторы позволят создавать вычислительные машины, которые будут обрабатывать информацию почти так же, как это делают современные электронные компьютеры, но значительно быстрее. В отдаленном же будущем оптические транзисторы сделают структуру вычислительных машин принципиально иной. В отличие от электронного транзистора, который в каждый момент времени работает лишь с одним сигналом, его оптический аналог может одновременно осуществлять большое число переключающих действий при использовании параллельных лазерных пучков. Более того, кристаллы, с которыми мы работаем в настоящее время, могут выполнять несколько последовательных переключений, сопровождающихся ступенчатым увеличением мощности на выходе при ступенчатом увеличении мощности падающего пучка, тогда как электронные транзисторы, которые используются в современных ЭВМ, имеют всего два уровня выходного сигнала. Применение устройств, имеющих более двух устойчивых состояний, может принципиально преобразовать машинную логику.

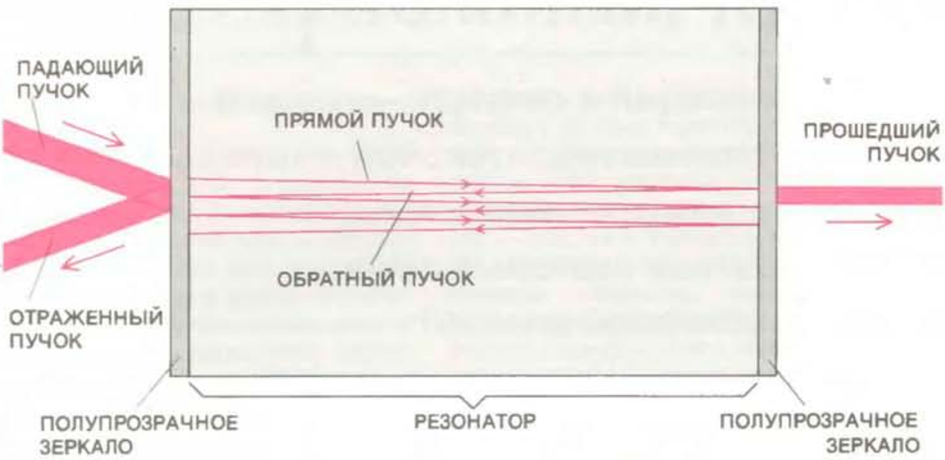
Для построения оптического компьютера помимо оптического транзистора нужны будут и другие схемные элементы. В нашей лаборатории в Университете Эрио-Уатта мы демонстрировали многие экспериментальные элементы, необходимые для оптических интегральных схем. Однако предстоит преодолеть немалые технические трудности, прежде чем такие схемы удастся освоить в промышленности. Тем не менее оптический компьютер является заманчивой перспективой относительно недалекого будущего.

Функции переключателя

В НАСТОЯЩЕЕ время все три основные функции вычислительной машины: выполнение арифметических действий и логических операций, а также запоминание информации — осуществляются с помощью устройств, имеющих два устойчивых состояния. При выполнении арифметических действий одно из этих состояний соответствует цифре 0, а второе — 1 в двоичной системе счисления. В логических процедурах одно из состояний переключателя принимается за истинное суждение, а другое — за ложное. Результаты арифметических и логических операций хранятся в памяти машины, элементы которой могут принимать одно из двух устойчивых состояний.

Используя двоичную алгебраическую логику, вычислительная машина определяет истинность суждений с помощью всего трех элементарных логических функций, которые именуют функцией И, функцией ИЛИ и функцией НЕ. При выполнении функции И утверждение является истинным, если все его составляющие элементы истинны. При выполнении функции ИЛИ утверждение считается истинным, если хотя бы один его элемент истинен. При выполнении функции НЕ истинное значение заключается в обратном утверждении. Эти три основные логические функции позволяют реализовать более сложные логические операции, а также арифметические действия, например сложение.

Таким образом, для создания вычислительной машины требуются устройства, способные в физической форме реализовать значения 0 и 1 или «истинно» и «ложно» и из которых можно строить более сложные блоки для



СХЕМАТИЧЕСКОЕ ИЗОБРАЖЕНИЕ ИНТЕРФЕРОМЕТРА ФАБРИ — ПЕРО, послужившего основой для оптических переключающих устройств, сконструированных авторами. Интерферометр имеет два плоских зеркала, расположенных параллельно и разделенных некоторым пространством. В это пространство, именуемое резонатором, помещается прозрачное вещество. Каждое зеркало частично отражает и частично пропускает падающий на него свет. Если пучок света направить на переднее зеркало, часть его отразится, а часть пройдет внутрь резонатора. Прошедший «прямой» пучок распространяется вдоль резонатора и падает на заднее зеркало. Часть света выходит из устройства, а часть отражается, образуя «обратный» пучок. Пропускание интерферометра зависит от длины резонатора, свойств заполняющего его вещества и длины световой волны. В оптическом переключателе вещество, заполняющее резонатор, обладает свойствами, позволяющими регулировать пропускание изменением интенсивности падающего пучка.

выполнения трех указанных логических функций. Очевидно, такое устройство должно иметь два четко различимых состояния. Чтобы создать быстродействующую машину, время переключения устройства из одного состояния в другое должно быть как можно меньше. В силу других соображений желательно, чтобы переключатель имел как можно меньшие размеры, был прост в производстве и потреблял минимальную мощность.

Первым электронным переключателем была вакуумная лампа — громоздкий и инерционный прибор, выделявший при работе значительное количество тепла, что сокращало срок его службы. В результате ламповые вычислительные машины имели большие физические размеры, но малую вычислительную мощность. В более мощных машинах приходилось непрерывно заменять выходящие из строя лампы.

Изобретение транзистора в 1947 г. позволило создать более компактный, быстрый и эффективный переключатель. В простейшем виде транзистор представляет собой трехслойную структуру из полупроводниковых материалов. Внешние слои в этой структуре называются эмиттером и коллектором, а средний слой — базой. Электрические свойства слоев подобраны таким образом, что незначительное изменение тока, протекающего от базы к коллектору, вызывает существенное изменение тока от эмиттера к коллектору.

Большой ток, протекающий через коллектор, можно принять за 1, а меньший ток — за 0. Транзисторы в комбинации с другими элементами схем позволяют создать структуры, выполняющие функции И, ИЛИ и НЕ. Такие структуры именуются логическими схемами. Так, схема, реализующая функцию И, имеет на выходе большой ток, если все входные сигналы велики. Схема ИЛИ дает большой ток, если хотя бы один из входных сигналов велик. Схема НЕ имеет на выходе большой ток при малом сигнале на входе и малый ток при большом значении входного сигнала.

Ограничения транзистора

В ЗАВИСИМОСТИ от размеров, материала, конструкции и мощности полупроводниковые логические схемы срабатывают за время от 1 нс до 1000 нс, то есть в идеальном случае самые быстродействующие схемы могут выполнять до миллиарда логических операций в секунду. Однако на практике быстродействие вычислительных машин значительно ниже. Руководствуясь физическими законами, которым подчиняются процессы, происходящие в полупроводниковых приборах, можно утверждать, что достигнутое к настоящему времени минимальное время переключения транзисторных схем порядка 1 нс близко к пределу.

Чтобы транзистор переключался из одного состояния в другое, необходимо, чтобы поток электронов пересек базу. Однако скорость передвижения электронов в полупроводнике имеет предел. Сократить время переключения можно за счет уменьшения толщины базы, но и ее можно уменьшить до определенного предела. По-видимому, этот предел уже достигнут, и поэтому всякие незначительные усовершенствования конструкции транзисторов уже не могут привести к существенному сокращению времени переключения; для этого требуется переход к принципиально новой технологии.

Одна их таких возможностей — отказ от электрического тока как средства передачи информации в цепях вычислительной системы и замена его сигналами иной природы. Наивысшей скоростью обладает свет (или электромагнитное излучение, если речь идет о других частях спектра), и поэтому он является наилучшим кандидатом на роль нового носителя информации. Имея это в виду, мы использовали когерентное излучение лазера для управления работой переключающих устройств, обладающих многими свойствами транзистора. Эти оптические ключи имеют два хорошо различимых выходных состояния; они могут служить элементами памяти и при соответствующем выборе материала и лазерных пучков позволяют строить простые и быстродействующие логические схемы.

При создании оптического транзистора была заимствована гениальная идея, заложенная в широко распространенном оптическом устройстве, известном под названием интерферометра Фабри — Перо, изобретенном французскими физиками Чарльзом Фабри и Альфредом Перо в 1896 г. Фабри и Перо использовали свой интерферометр для измерения длин волн различных цветов света; впоследствии этот прибор широко использовался для самых различных целей. Простейший интерферометр Фабри — Перо состоит из двух плоских зеркал, расположенных параллельно друг другу и разделенных некоторым пространством. В это пространство, именуемое полостью или резонатором, помещается вещество, пропускающее свет определенной длины волны.

Каждое из зеркал частично отражает и частично пропускает падающий на него свет. Такие частично отражающие зеркала, именуемые обычно полупрозрачными, встречаются повсюду. Обычное стекло витрины, например, яляется таким зеркалом: оно отражает свет и дает изображение стоящего перед ним наблюдателя, и в то же время оно пропускает свет, позволяя видеть товары, находящиеся за стеклом. Отчетливость, с которой мы видим свое отражение в стекле и предметы за ним, зависит от того, в какой пропорции падающий свет разделяется на отраженный и прошедший.

Забудем на время о веществе, запо-

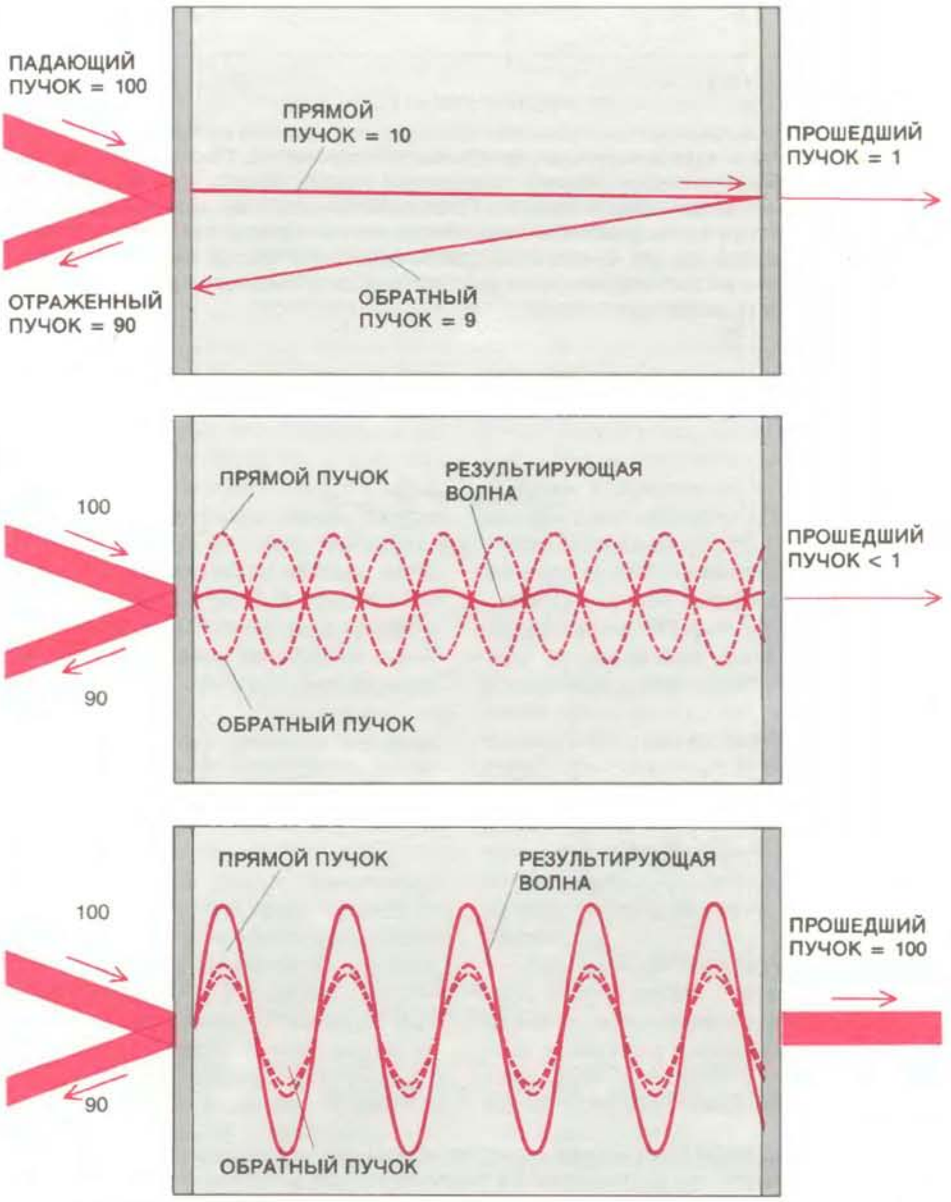
лняющем резонатор интерферометра, и рассмотрим, что произойдет, когда пучок света падает на зеркало, образующее переднюю стенку интерферометра. Допустим, зеркало отражает 90 % падающего света и пропускает 10 % (такое соотношение весьма близко к условиям нашего эксперимента), поэтому лишь десятая часть световой энергии проникает внутрь интерферометра и распространяется вдоль резонатора к заднему зеркалу, образуя «прямой» пучок. Заднее зеркало обладает теми же свойствами, что и переднее. Поэтому 90 % прямого пучка отразится от заднего зеркала обратно в резонатор, образуя «обратный» пучок, а 10 % выйдет из интерферометра. Так как интенсивность прямого пучка в десять раз слабее входного падающего пучка, то выходной пучок, прошедший через интерферометр, будет иметь интенсивность, в 100 раз меньшую по сравнению с интенсивностью входного пучка. Понятно, что обратный пучок, двигаясь в обратном направлении, дойдет до внутренней поверхности переднего зеркала, где снова разделится на прошедшую и отраженную части. Ослабленный пучок, многократно отражаясь от зеркал резонатора, будет становиться все слабее и слабее до тех пор, пока весь свет не выйдет из интерферометра. Каждый проход пучка вносит определенный вклад в выходной пучок, однако, за исключением первого, эти вклады чрезвычайно малы: можно считать, что через интерферометр проходит лишь сотая часть входного пучка.

Интерференция в резонаторе

ЕСЛИ БЫ физические процессы в интерферометре исчерпывались теми явлениями, о которых шла речь, то это устройство не представляло бы интереса в качестве переключателя для оптического компьютера. Интенсивность выходного пучка в данном случае можно увеличить только путем изменения свойств зеркал. Однако мы не учли одного обстоятельства. В действительности прямой и обратный пучки нельзя разделить так, как это сделано в нашем схематическом описании. В настоящем интерферометре взаимодействие прямых и обратных пучков влияет на интенсивность светового поля в резонаторе, которое в свою очередь определяет пропускание интерферометра.

Хорошо известно, что свет обладает свойствами как волны, так и частицы. В одних случаях удобнее пользоваться описанием света как волны, а в других — как частицы. В данном случае удобнее воспользоваться волновым описанием. Прямой и обратный пучки совместно с многократно отраженными и ослабленными пучками можно рассматривать как волны, распространяющиеся в резонаторе. Входной пучок падает на интерферометр перпендикулярно его поверхности, и поэтому все пучки в резонаторе распространяются по одному и тому же пути.

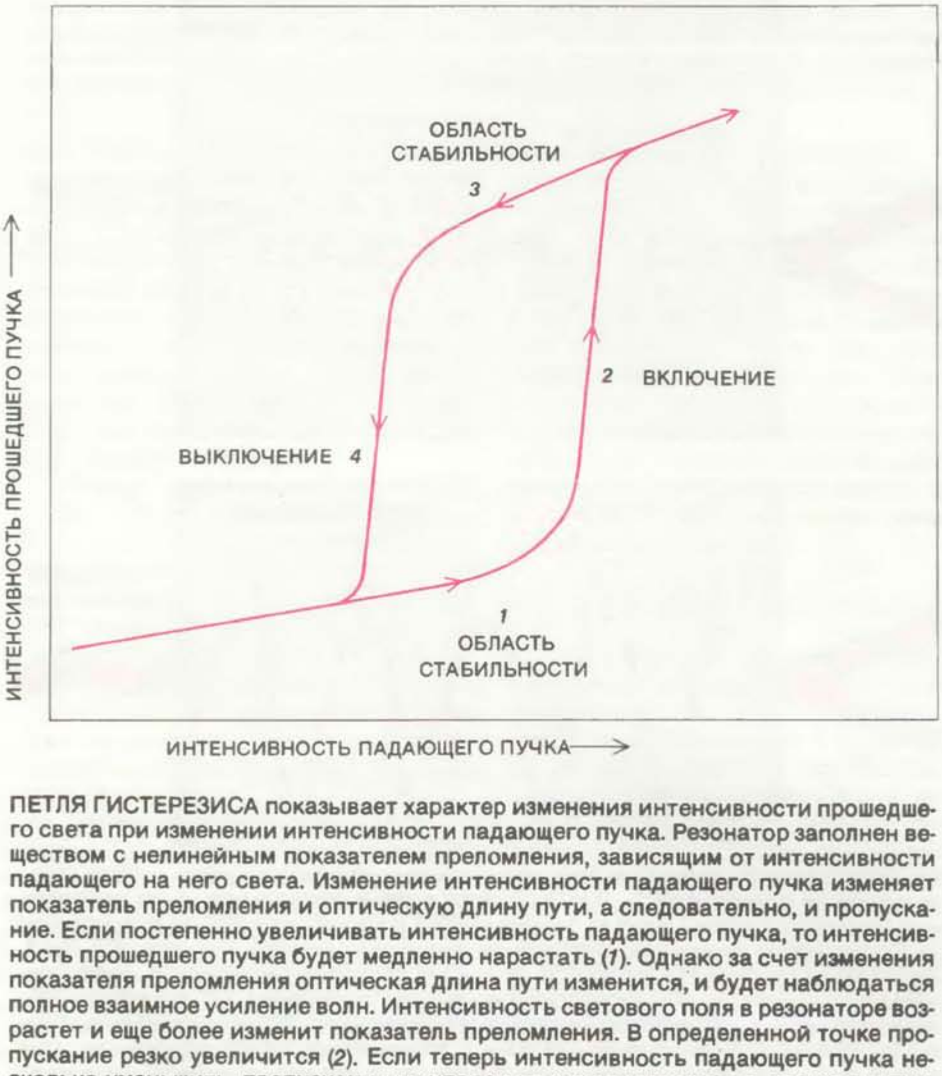
Помня о том, что в резонаторе имеется много отраженных пучков, мы все же можем рассматривать взаимодействие только двух из них: прямого и обратного. Результат взаимодействия зависит от соотношения фаз волн, то есть от относительного положения их вершин и впадин. Если вершины одной волны совпадают с вершинами другой волны, а впадины — с впадинами, то в резонаторе наблюдается конструктивная интерференция (с взаимным усилением волн). И наоборот, если вершины одной волны совпадают с впадинами другой, то наблюда-



ИНТЕРФЕРЕНЦИЯ прямого и обратного пучков в резонаторе Фабри — Перо сильно влияет на интенсивность светового поля в нем и, следовательно, на пропускание. Переднее и заднее зеркала отражают 90% и пропускают 10% падающего света. Соотношение интенсивностей пучков показано на *верхнем* рисунке. 10% падающего пучка проникает в резонатор, достигает поверхности заднего зеркала и выходит из резонатора. Интенсивность на выходе составляет 1% интенсивности входного пучка, а интенсивность обратного пучка равна 9% входного пучка. Прямой и обратный пучки взаимодействуют, образуя результирующую волну. Она может быть ослабленной при интерференции с взаимным ослаблением волн (*средний рисунок*). Интенсивность света в резонаторе при этом мала и пропускание составляет менее 1%. Если обе волны усиливают друг друга, то результирующая волна велика (*нижний рисунок*). Так как в резонаторе распространяется много прямых и обратных волн, то интенсивность светового поля в резонаторе может в 10 раз превышать интенсивность падающего пучка. Поэтому при интерференции с взаимным усилением волн прошедший пучок по интенсивности равен исходному падающему пучку.



ФУНКЦИЯ ЭЙРИ выражает соотношение между пропусканием интерферометра и оптической длиной пути в веществе, заполняющем резонатор. Пропускание измеряется отношением интенсивностей падающего и прошедшего пучков. Максимальное значение пропускания равно 1. Пропускание максимально, когда фазы пучков в резонаторе совпадают. Оптическая длина пути равна толщине слоя вещества, умноженной на его показатель преломления. Интерференция с полным взаимным усилением наблюдается, когда оптическая длина пути кратна половине длины волны (λ) падающего света.



ПЕТЛЯ ГИСТЕРЕЗИСА показывает характер изменения интенсивности прошедшего света при изменении интенсивности падающего пучка. Резонатор заполнен веществом с нелинейным показателем преломления, зависящим от интенсивности падающего на него света. Изменение интенсивности падающего пучка изменяет показатель преломления и оптическую длину пути, а следовательно, и пропускание. Если постепенно увеличивать интенсивность падающего пучка, то интенсивность прошедшего пучка будет медленно нарастать (*1*). Однако за счет изменения показателя преломления оптическая длина пути изменится, и будет наблюдаться полное взаимное усиление волн. Интенсивность светового поля в резонаторе возрастет и еще более изменит показатель преломления. В определенной точке пропускание резко увеличится (*2*). Если теперь интенсивность падающего пучка несколько уменьшить, пропускание не упадет до исходного уровня, так как интенсивность света в резонаторе еще достаточно велика и показатель преломления и оптическая длина пути соответствуют максимуму пропускания; пропускание падает медленно (*3*). При дальнейшем снижении интенсивности падающего пучка произойдет значительная потеря интенсивности света в резонаторе и пропускание резко упадет до начального уровня (*4*). В петле гистерезиса участки *1* и *3* представляют собой области стабильности, используемые в работе оптического переключающего устройства.

ется деструктивная интерференция (с взаимным ослаблением волн). Между этими двумя экстремальными явлениями, безусловно, может наблюдаться и любая промежуточная картина.

При взаимодействии двух волн в результате их интерференции возникает третья волна. Амплитуда результирующей волны в каждой точке резонатора равна сумме амплитуд составляющих ее волн в этой точке. Если волны взаимно усиливают друг друга, эта амплитуда велика, если же они ослабляют друг друга, она мала. Интерференция света оказывает совершенно неожиданное действие на пропускание интерферометра Фабри — Перо. При полном взаимном ослаблении волн интенсивность светового поля в резонаторе близка к нулю и пропускание незначительно. При полном взаимном усилении многочисленные прямые и обратные пучки усиливают друг друга и создают в резонаторе свет, интенсивность которого может в 10 раз превышать интенсивность падающего пучка. Так как заднее зеркало пропускает 10 % падающего на него света, при полном взаимном усилении интенсивность пучка на выходе интерферометра равна интенсивности пучка на его входе.

Влияние показателя преломления

Фазы прямого и обратного пучков, а следовательно, и пропускание интерферометра можно регулировать изменением расстояния между зеркалами и длиной волны падающего света. Однако ни один из этих методов не пригоден для оптического транзистора. Более важным является то обстоятельство, что соотношение фаз можно изменять подбором оптических свойств вещества, заполняющего резонатор, о чем мы пока не говорили.

Свет распространяется с наибольшей скоростью в вакууме. В любой другой прозрачной среде скорость света меньше, и она зависит от свойств среды. Именно замедление световой волны создает известное явление преломления; показатель преломления вещества численно равен отношению скорости света в вакууме к скорости света в веществе. Например, если показатель преломления определенного сорта стекла равен 1,5, то это означает, что свет в этом стекле распространяется со скоростью не 300 000 км/с, как в вакууме, а 200 000 км/с. Замедление прохождения света сопровождается уменьшением длины волны, однако частота колебаний при этом не меняется. Отсюда понятно, что, даже не изменяя длину резонатора или длину волны падающего света, можно изменить фазовые соотношения распро-



ФОРМА ПЕТЛИ ГИСТЕРЕЗИСА оптически бистабильного устройства зависит от длины интерферометра, длины волны падающего света и свойств вещества, заполняющего резонатор. В обычной петле гистерезиса стабильные области протяженны и петля широкая (*слева*). Каждому значению интенсивности падающего пучка могут соответствовать одно или два значения интенсивности прошедшего пучка. Конструкцию интерферометра можно изменить так, что стабильные области станут короче, а петля сузится (*в середине*). Можно получить и однозначную кривую, в которой каждому значению интенсивности падающего пучка соответствует только одно значение интенсивности прошедшего пучка (*справа*). Кривая аналогична характеристике тока коллектора транзистора, и интерферометр с такой кривой пропускания используется как оптический транзистор.

страняющихся в нем пучков только с помощью выбора вещества, заполняющего резонатор, так как за счет эффекта преломления изменится длина световой волны в резонаторе. При этом в зависимости от значения показателя преломления выбранного вещества можно получить либо взаимное усиление, либо взаимное ослабление прямого и обратного пучков света.

В анализе процессов в оптических системах важным является показатель, называемый оптической длиной пути, или просто оптической длиной. Для среды, заполняющей резонатор интерферометра, этот показатель равен физической длине резонатора, умноженной на показатель преломления среды. Можно показать, что полная конструктивная интерференция возникает, когда оптическая длина резонатора равна целому числу полуволн падающего света. Очевидно, что взаимное усиление световых волн может происходить при многих различных значениях оптической длины. Если же значение оптической длины лежит посредине между двумя расстояниями, кратными целым числам полуволн (т. е. отличается от обоих расстояний на четверть длины волны), то наблюдается полная деструктивная интерференция.

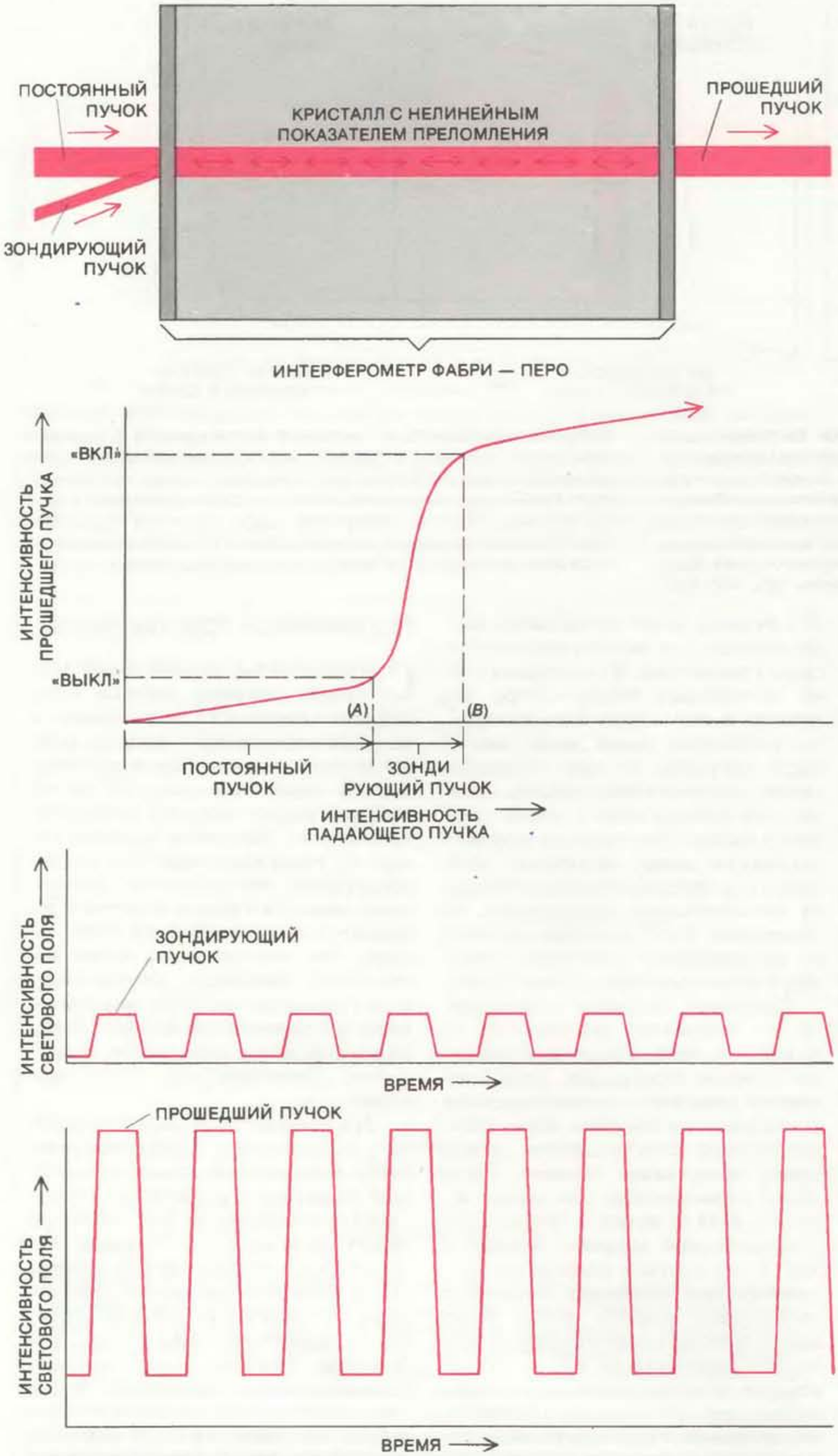
Так как интенсивность света, прошедшего через интерферометр, изменяется в зависимости от фазовых соотношений между пучками, то из этого следует, что интенсивность прошедшего света зависит от оптической длины. Функция, которая описывает соотношение между оптической длиной и пропусканием интерферометра (равное отношению интенсивности пучка на выходе к интенсивности падающего пучка), называется функцией Эйри. Эта функция имеет чрезвычайно важное значение для конструкции оптического транзистора. В создаваемых нами устройствах Фабри — Перо эта функция имеет острые пики в точках, где оптическая длина равна целому числу полуволн; во всех остальных точках, лежащих между пиками, значение этой функции мало, и кривая изменяется плавно. Поэтому если изменять оптическую длину, постепенно приближая ее к значению, соответствующему максимальному пропусканию, то пропускание будет меняться медленно до определенного порогового значения, а затем произойдет резкий скачок.

Поскольку оптическая длина зависит от показателя преломления, то можно получить любое фиксированное значение пропускания, введя в резонатор вещество с соответствующим показателем преломления. Более того, заменяя одно вещество другим, можно менять пропускание. Конечно, такой способ переключения, как замена вещества, вряд ли является удобным для вычислительной машины. Однако в XIX в. его считали единственно возможным для изменения показателя преломления. Джеймс Кларк Максвелл в своей фундаментальной работе по электромагнитному излучению утверждал, что преломление и другие явления взаимодействия света с веществом не зависят от интенсивности света. Если бы это было справедливо во всех случаях, показатель преломления был бы константой, а интенсивность света на выходе интерферометра была бы пропорциональна интенсивности падающего пучка. В определенных условиях многие вещества проявляют именно такие свойства, и в этом случае говорят, что они имеют линейный показатель преломления.

Нелинейное преломление

С ПОЯВЛЕНИЕМ лазеров стало возможным получать мощные когерентные излучения (т. е. излучения, в которых все вершины и впадины волн согласованы между собой во времени), и тогда стало очевидным, что не все вещества имеют линейный показатель преломления. Направляя излучение лазера на различные вещества, ученые обнаружили, что показатель преломления некоторых из них меняется в зависимости от интенсивности пучка; говорят, что такие вещества имеют нелинейный показатель преломления. Если такое вещество поместить в резонатор интерферометра Фабри —Перо, то его показатель преломления можно менять интенсивностью падающего пучка.

Это явление заслуживает подробного рассмотрения. Представим резонатор определенной длины, заполненный веществом с нелинейным показателем преломления; на этот резонатор падает пучок лазерного излучения, интенсивность которого можно менять. Пусть вначале интенсивность пучка такова, что оптическая длина резонатора, обусловленная показателем преломления вещества, имеет значение, лежащее между пиками кривой Эйри, и пропускание в этом случае мало. Если теперь интенсивность пучка медленно увеличивать, то показатель преломления и оптическая длина будут постепенно меняться, а пропускание будет медленно возрастать. В определенный момент между показателем преломления и силой света образуется положительная обратная связь — они начнут увеличивать друг друга. За счет изменения показателя преломления световые пучки будут смещаться все ближе



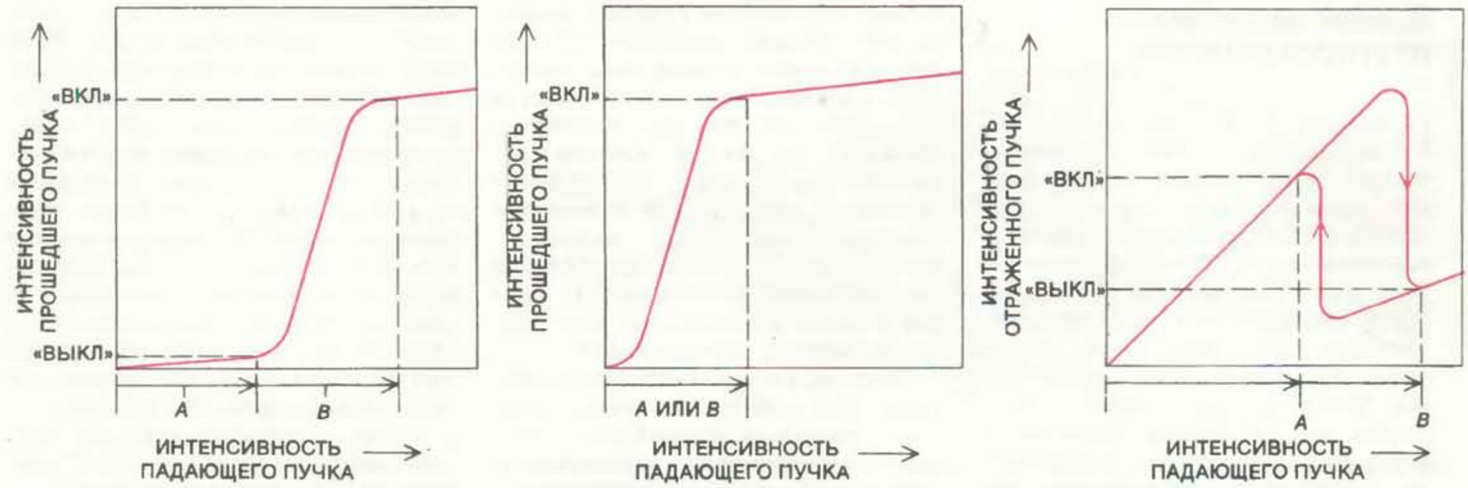
ОПТИЧЕСКИЙ ТРАНЗИСТОР, сконструированный авторами статьи, работает от двух лазерных пучков, фокусируемых на переднем зеркале. Резонатор заполнен веществом с нелинейным показателем преломления (*вверху*). «Постоянный» пучок имеет большую и неизменную интенсивность. «Зондирующий» — слабее и может изменяться. Интенсивность постоянного пучка (*A*) соответствует малому пропусканию и лежит ниже крутого участка кривой пропускания (*в середине*). Суммарная интенсивность двух пучков (*B*) резко повышает пропускание. Нижний стабильный участок (выключенное состояние) может соответствовать 0 в двоичной системе счисления, а верхний (включенное состояние) — 1. На крутом участке кривой малые изменения интенсивности зондирующего пучка вызывают большие изменения интенсивности прошедшего пучка (*внизу*).

к полному взаимному усилению, и интенсивность светового поля в резонаторе будет расти, еще более изменяя показатель преломления. При некотором критическом уровне интенсивности падающего пучка эффект обратной связи становится чрезвычайно сильным. На кривой Эйри это будет соответствовать переходу на самый крутой участок, где пропускание резко возрастает.

В пике функции Эйри отношение интенсивностей прошедшего и падающего пучков равно 1. Если теперь интенсивность лазерного пучка несколько уменьшить, пропускание резонатора не упадет до первоначального низкого уровня; интенсивность светового поля в резонаторе достаточно велика, чтобы значения показателя преломления и оптической длины удерживались на уровне, соответствующем максимуму пропускания. Если интенсивность пучка снизить еще больше, то пропускание плавно уменьшится. В определенный момент, однако, показатель преломления и интенсивность светового поля в резонаторе начнут ослаблять друг друга, и тогда даже небольшое изменение интенсивности падающего пучка настолько изменит показатель преломления, что значительно повлияет на интенсивность светового поля в резонаторе. Пропускание внезапно и резко снизится.

Если построить график зависимости интенсивности излучения на выходе от интенсивности падающего пучка, то получится кривая в виде петли, состоящей из четырех частей (нижний рис. на с. 18). От точки с минимальным значением интенсивности падающего пучка пропускание постепенно увеличивается: в этой части кривой даже большое изменение интенсивности падающего пучка лишь незначительно увеличивает пропускание. При пороговом значении интенсивности кривая круто устремляется вверх: в этой части небольшое изменение интенсивности падающего пучка приводит к резкому изменению пропускания. Возвращение к исходной точке идет по другому участку кривой. Сначала пропускание падает медленно и остается высоким, даже если интенсивность падающего пучка уменьшается на значительную величину. Когда же интенсивность попадет в область второго порогового значения, кривая круто обрывается вниз: при этом небольшое снижение интенсивности ведет к значительному ослаблению пропускания.

Такую кривую называют петлей гистерезиса, а оптическое устройство, изменяющее свои свойства по такому циклу, называют оптически бистабильным, так как оно имеет две стабильные области, в которых интенсивность светового пучка на выходе изменяется мало при изменении интенсив-



ЛОГИЧЕСКИЕ ЭЛЕМЕНТЫ можно строить из оптических транзисторов. Из логических элементов формируются блоки вычислительных машин, выполняющие логические и арифметические операции. Для этого требуются три типа элементов. Элемент И имеет сигнал на выходе только в том случае, если есть сигналы на всех его входах (*слева*). Если имеются два лазерных пучка и под действием любого из них пропускание лежит несколько ниже крутого участка кривой, а при действии обоих пучков рабочая точка переходит на крутой участок, то реализуется операция И. Элемент ИЛИ дает большой сигнал на выходе, если на любом из его входов имеется высокий уровень сигнала (*в середине*). Если при действии любого из двух пучков значение пропускания лежит вверху крутого участка кривой, то реализуется операция ИЛИ. Элемент НЕ имеет высокий сигнал на выходе, когда входной сигнал мал, и наоборот. Оптический элемент НЕ работает на отраженном пучке света (*справа*), за счет которого создается петля гистерезиса; при высоком сигнале на входе (*B*) сигнал на выходе слабый, а низкий входной сигнал (*A*) создает большой сигнал на выходе.

ности падающего пучка. Следует отметить, что в бистабильном устройстве определенному значению интенсивности падающего пучка могут соответствовать два уровня интенсивности выходного пучка: иначе говоря, пропускание зависит от предыдущего состояния резонатора.

Бистабильные оптические устройства

Существование оптической бистабильности было предсказано еще в 1969 г. Э. Жеке и его коллегами из Массачусетского технологического института. Впервые ее наблюдали сотрудники фирмы Bell Laboratories X. Джиббс, С. Мак-Колл и Т. Венкатесан в 1976 г. С тех пор бистабильные оптические устройства стали предметом исследований во многих лабораториях.

Для изготовления оптически бистабильного устройства важно найти вещество, обладающее значительной нелинейностью показателя преломления. В нашей работе мы использовали сложный полупроводник — антимонид индия, обладающий чрезвычайно нелинейной характеристикой. Как и большинство полупроводников, антимонид индия непрозрачен для видимой части спектра, но прозрачен для некоторой области инфракрасного излучения. В качестве источника мы использовали лазер на окиси углерода, длину волны которого можно регулировать в узком диапазоне инфракрасной части спектра.

В качестве инерферометра мы использовали небольшой прямоугольный кристалл антимонида индия. Обычно каждая грань кристалла равна нескольким миллиметрам, но мы делали устройства и на кристаллах значительно меньших размеров. В некоторых случаях передние и задние грани кристаллов покрывались отражающим веществом, но это покрытие необязательно: полированная поверхность кристалла сама может служить зеркалом. Позади заднего зеркала помещался прибор для измерения интенсивности прошедшего пучка.

С точки зрения атомного строения вещества можно по-разному объяснить изменение показателя преломления полупроводника. Механизм, который действует в антимониде индия, достаточно сложен и до сих пор еще до конца не изучен. Однако некоторое объяснение этого процесса имеется, и мы изложим его в упрощенном виде.

Взаимодействие электромагнитного излучения с твердым веществом происходит в основном за счет возбуждения электронов вещества. Если электромагнитное излучение лежит в видимой или инфракрасной части спектра, то во взаимодействии участвуют главным образом электроны внешних оболочек атомов. В хорошем проводнике, например в металле, электроны с внешних оболочек могут свободно передвигаться в объеме вещества; в хорошем изоляторе каждый электрон прочно связан с определенным атомом, и, чтобы сместить его, требуется значительная энергия. В полупроводнике наблюдается промежуточное явление: каждый наружный электрон связан с атомом, но достаточно приложить небольшую энергию, и эти электроны станут свободными.

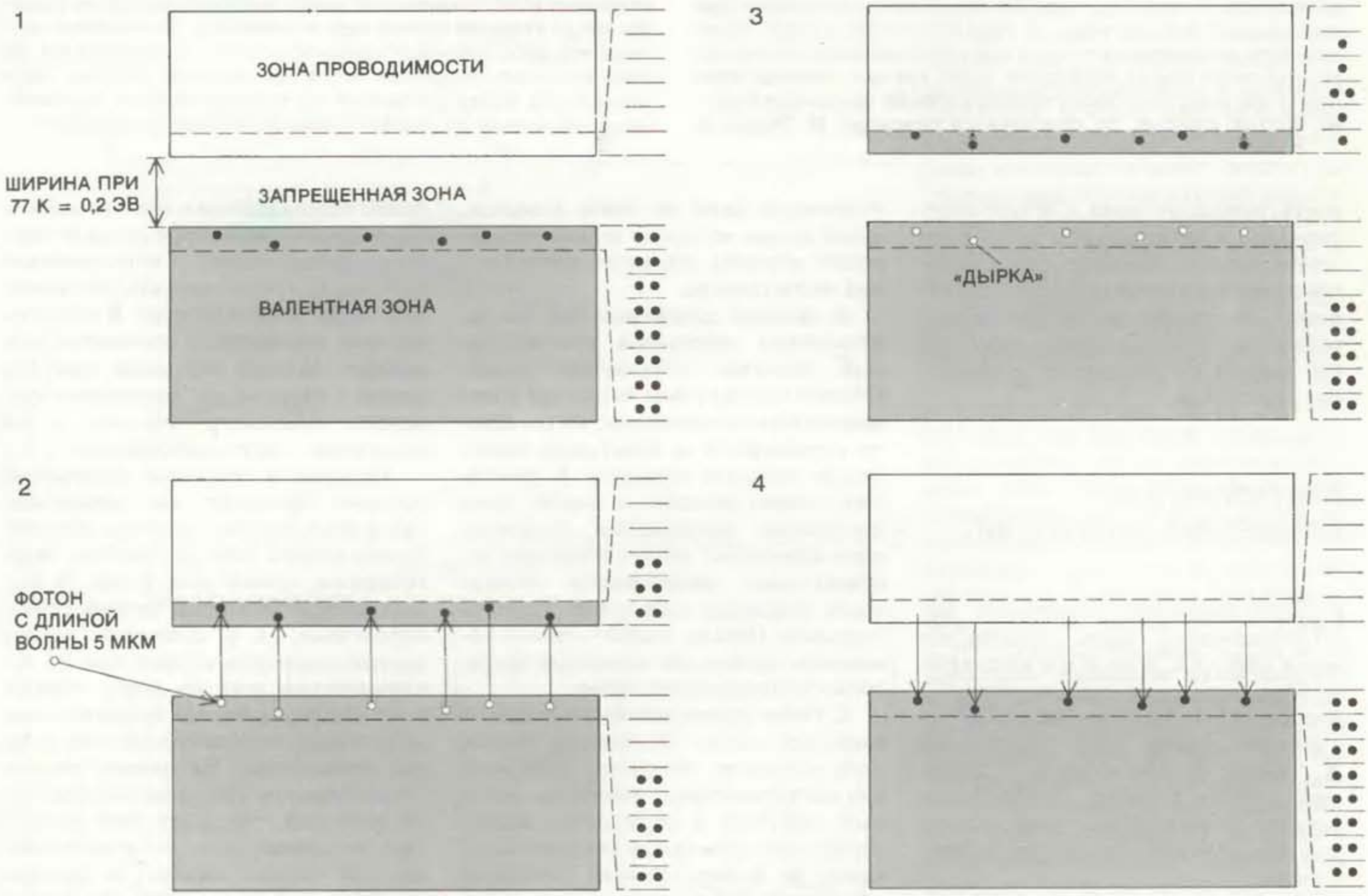
Различие в поведении электронов, которое определяет как оптические, так и электрические свойства веществ, можно понять, если рассмотреть энергетические уровни электронов. В изолированном атоме эти уровни строго определены, а в сплошных средах уровни различных атомов как бы накладываются друг на друга, образуя широкие разрешенные энергетические зоны и расположенные между ними запрещенные зоны. На уровнях энергии запрещенных зон не может находиться ни один электрон. Если зона полностью зополнена, т. е. все принадлежащие ей уровни заняты, то каждый электрон этой зоны «привязан» к определенному атому. Можно дать следующее объяснение этому явлению: для приведения электрона в движение необходима небольшая дополнительная энергия, а в заполненной зоне нет ни одного свободного уровня, который мог бы занять электрон, получивший дополнительную энергию. Если зона занята не полностью, то возбужденный электрон может легко перейти на более высокий уровень этой зоны, т. е. может стать свободным.

Зонная структура полупроводников

В МЕТАЛЛАХ разрешенные зоны заполнены не полностью. В изоляторе зона самых высоких уровней энергии (валентная зона) заполнена полностью и отделена широкой запрещенной зоной от следующей с более высокими энергетическими уровнями зоны (зоны проводимости), которая пуста. Электрон может покинуть свою зону в изоляторе только в том случае, если ему будет сообщена энергия, достаточная для преодоления запрещенной зоны, но такое событие маловероятно. У полупроводника валентная зона тоже заполнена, а зона проводимости пуста, но промежуток между этими зонами у́же. Существуют различные способы, как сообщить электрону в полупроводнике то довольно малое количество энергии, которое ему необходимо для перехода с одного из наивысших уровней валентной зоны на один из нижних уровней зоны проводимости. Когда электрон переходит в зону проводимости, он оставляет в валентной зоне «дырку», которая в некотором смысле ведет себя как положительный заряд, равный по величине отрицательному заряду электрона. Когда этот электрон возвращается в валентную зону, он соединяется с дыркой, и дырка уничтожается; этот процесс называется рекомбинацией.

Энергия, которая необходима электрону для пересечения запрещенной зоны, зависит от температуры. Большинство наших работ выполнено на антимониде индия при температуре 77 К. При этой температуре ширина запрещенной зоны антимонида индия составляет около 0,2 эВ (1 эВ — это энергия, которую приобретает электрон при ускорении разностью потенциалов 1 В). Большинство других полупроводников имеет гораздо более широкие запрещенные зоны. Нелинейность показателя преломления возрастает обратно пропорционально квадрату ширины запрещенной зоны. Благодаря малой ширине запрещенной зоны нелинейность антимонида индия от 100 до 1000 раз превосходит нелинейность арсенида галлия — другого полупроводника, который использовался в бистабильных оптических устройствах. И хотя желательно, чтобы запрещенная зона была неширокой, при слишком узкой зоне оптический переключатель работать не может.

Один из способов, которым электрону можно сообщить энергию, необходимую для перехода в зону проводимости, заключается в облучении вещества. Чтобы рассмотреть этот процесс, удобно использовать описание света как потока частиц. Частица, или квант, света называется фотоном. Энергия фотона обратно пропорцио-

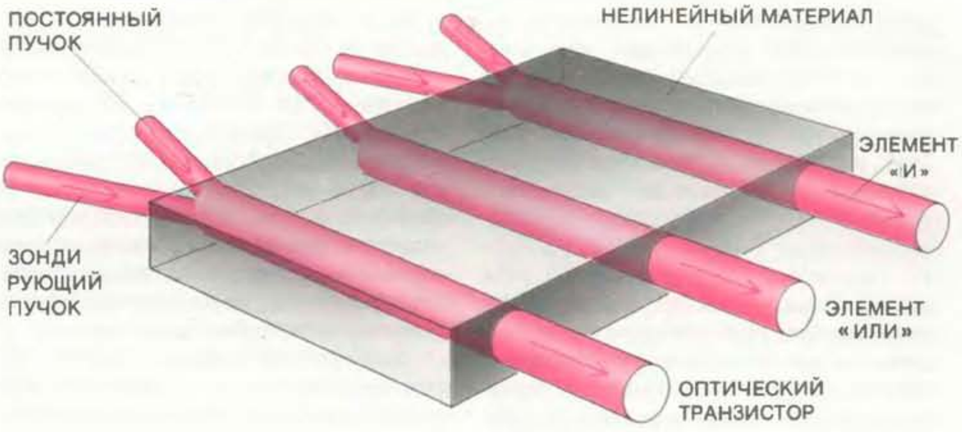


ПОКАЗАТЕЛЬ ПРЕЛОМЛЕНИЯ антимонида индия изменяется при изменении интенсивности падающего на него излучения. Антимонид индия — полупроводник, который авторы использовали для оптического транзистора. Уровни энергий, которые может занимать электрон в кристалле полупроводника, сгруппированы в зоны; в зоне большое количество близко расположенных уровней, и на каждом из них могут находиться только два электрона. Между разрешенными зонами расположена запрещенная зона. При низкой температуре нижняя зона (валентная) полностью заполнена, а верхняя (зона проводимости) — пуста (*1*). Для перехода из валентной зоны в зону проводимости электрон должен возбудиться, то есть приобрести дополнительную энергию, эквивалентную ширине запрещенной зоны; в антимониде индия эта энергия составляет около 0,2 эВ. Электрон может получить эту энергию от фотона с длиной волны 5 мкм (инфракрасная область спектра). Фотоны с такой длиной волны генерируются лазером на окиси углерода (*2*). Когда электрон переходит в зону проводимости, в валентной зоне образуется «дырка» с положительным зарядом. Тепловая и вибрационная энергии рассеивают электроны на нижних уровнях зоны проводимости (*3*). Когда они заполнены, фотоны перестают поглощаться и показатель преломления кристалла изменяется. Когда луч выключают, электроны возвращаются в валентную зону и рекомбинируют с дырками (*4*). Показатель преломления становится прежним.

нальна длине волны. Поглощенный кристаллом фотон передает свою энергию электрону. Если фотон имеет достаточную энергию, электрон переходит в зону проводимости. Длина волны, эквивалентная энергии 0,2 эВ, равна приблизительно 5 мкм, то есть она лежит в инфракрасной части спектра и может генерироваться лазером на окиси углерода.

Если облучить антимонид индия лазером, настроенным на указанную длину волны, большое количество электронов перейдет в зону проводимости. Как только электроны достигнут нижних уровней зоны проводимости, тепловая энергия кристаллической решетки рассеет их по близлежащим энергетическим уровням. В результате возникает случайное распределение электронов по уровням нижней части зоны проводимости.

Именно переход электронов и последующее их рассеяние обусловливают изменение показателя преломления антимонида индия. В квантомеханическом понимании взаимодействия света с веществом преломление и поглощение тесно сязаны между собой. В самом деле, одно всегда влечет за собой другое: если пучок излучения поглощается, он также преломляется, и наоборот, хотя оба эти явления могут проявляться в разной степени для различных длин волн одного пучка. Создавая условия, препятствующие поглощению, мы в то же время препятствуем и преломлению и тем самым изменяем показатель преломления. Как воспрепятствовать поглощению? Существует предел количества электронов, которые можно перевести на нижние уровни зоны проводимости. Если этот предел будет достигнут, ни один дополнительный фотон с энергией, достаточной для такого перехода, не будет поглощен; в этом случае говорят, что поглощение фотонов достигло насыщения. Насыщение поглощения препятствует сопутствующему преломлению, и показатель преломления изменяется. Очень скоро после выключения лазерного пучка перешедшие электроны возвращаются в валентную зону и рекомбинируют с дырками. Поглощение снова становится возможным, и показатель преломления возвращается к своему исходному значению.



ПАРАЛЛЕЛЬНАЯ ОБРАБОТКА — новая возможность, открываемая оптическими переключателями. Она может привести к изменению конструкции и производительности вычислительных машин. На одном кристалле нелинейного вещества можно сфокусировать несколько лазерных пучков, которые будут проходить через кристалл каждый по своему пути. Каждый луч можно использовать для выполнения разных операций. Например, при наличии трех лучей на одном из них можно выполнять операцию И, на втором — операцию ИЛИ, а третий может работать как транзистор. Если три пучка получить расщеплением одного исходного пучка, можно выполнить сразу три операции с одним сигналом (что невозможно в ЭВМ). Такая возможность потребует новой организации процесса обработки информации.





МНОГОКРАТНАЯ БИСТАБИЛЬНОСТЬ — свойство оптических переключателей, которое может привести к разработке новой системы машинной логики. В интерферометре существует много значений оптической длины пути, которые равны целому числу полуволн излучения. Каждому такому значению соответствует пик в функции Эйри и свой максимум пропускания. При постепенном увеличении интенсивности падающего пучка оптическая длина нелинейного кристалла в резонаторе принимает ряд значений, соответствующих максимумам пропускания. Поэтому зависимость интенсивности пропускания от интенсивности падающего пучка представляется в виде кривой с несколькими вертикальными ступеньками, каждой из которых соответствует пик на кривой Эйри (*слева*). Если интенсивность постоянного пучка в оптическом транзисторе увеличивать от ступеньки к ступеньке, то, используя один и тот же слабый зондирующий пучок, можно получать при каждом последовательном переключении все более мощный выходной пучок. Так как высота вертикальных ступенек на кривой пропускания (*слева*) увеличивается, то коэффициент усиления, равный отношению интенсивности прошедшего пучка к интенсивности зондирующего пучка будет нарастать при каждом переключении (*справа*). Различные уровни пропускания можно использовать для представления значений в двоичной системе.

Логические схемы из оптических элементов

НЕОБЫЧНЫЕ оптические свойства антимонида индия, о которых шла речь, оказались пригодными для изготовления элементов оптического компьютера. Очевидно, что гистерезис можно использовать в элементах памяти. Устойчивое состояние интерферометра на высоком уровне пропускания можно принять в качестве одного логического или арифметического значения, а устойчивое состояние на низком уровне пропускания — для представления другого значения в двоичной системе. Любое из этих состояний может поддерживаться сколь угодно долго с помощью пучка промежуточной интенсивности.

Для оптического переключателя в блоках обработки информации, а не памяти гистерезис не нужен. Слегка изменив оптическую схему, петлю гистерезиса можно сузить и даже совсем устранить ее. В результате получим однозначную кривую пропускания, у которой каждому значению интенсивности падающего пучка будет соответствовать лишь одно значение пропускания. Можно также изменить и характер этой кривой. Наиболее подходящей является кривая, у которой вначале имеется пологий участок, где малым значениям интенсивности падающего пучка соответствуют малые и почти постоянные значения пропускания,

затем крутой участок, на котором незначительным изменениям интенсивности пучка на входе соответствуют большие изменения пропускания, и, наконец, третий такой же пологий участок, как и первый, но в зоне бо́льших значений интенсивности падающего пучка. Ее форма напоминает характеристику тока коллектора электронного транзистора. Кривая такого вида принимается за основу в оптическом транзисторе. Напомним, что это устройство мы называем трансфазором, так как его работа основана на управлении фазой световой волны, распространяющейся внутри него.

В трансфазоре два хорошо отъюстированных лазерных пучка фокусируются в одной точке на передней грани кристалла нелинейного вещества, такого, как антимонид индия. Один из пучков, именуемый постоянным, имеет относительно большую и неизменную интенсивность. Второй пучок, именуемый зондирующим, имеет значительно меньшую интенсивность и может модулироваться. Интенсивность постоянного пучка подобрана таким образом, что вызванное им пропускание лежит немного ниже крутого участка кривой пропускания. Интенсивность зондирующего пучка такова, что, когда он добавляется к постоянному пучку, рабочая точка переходит на участок кривой, где пропускание максимальное.

Так как кривая пропускания трансфазора очень крута, то даже незначительная модуляция зондирующего пучка вызывает сильное увеличение пропускания прибора. Аналогию трансфазора с электронным транзистором здесь можно продолжить. Постоянный пучок аналогичен постоянному току смещения, который протекает от эмиттера к коллектору транзистора. Зондирующий пучок аналогичен меньшему току, протекающему от базы к коллектору. Как небольшое изменение тока базы вызывает в транзисторе увеличение тока коллектора, так и небольшое изменение зондирующего пучка вызывает в трансфазоре бо́льшую проводимость.

Подобно транзистору, трансфазор может переключаться в одно из двух четко различимых состояний. Это переключение происходит чрезвычайно быстро. Хотя нам и не удалось непосредственно измерить время переключения, косвенные доказательства говорят о том, что оно равно нескольким пикосекундам.

Когда оптический транзистор создан, за этим должна следовать сборка прототипа логической схемы. Один и тот же трансфазор может служить как элементом И, так и элементом ИЛИ в зависимости от подведенных к нему пучков. Если два падающих пучка подобраны так, что ни один из них сам по себе не способен переключить трансфазор, а оба вместе обладают достаточной интенсивностью для его переключения, то образуется оптический элемент И. Пропускание велико лишь тогда, когда на трансфазор направлены оба пучка. Если падающие пучки подобраны так, что любой из них может переключить трансфазор, то образуется оптический элемент ИЛИ. Пропускание велико, если трансфазор облучается хотя бы одним пучком.

Элемент НЕ можно создать, используя в качестве выходного сигнала отраженный пучок. Так как он является инверсией прошедшего пучка, то повышение интенсивности падающего пучка приводит к снижению интенсивности на выходе, и наоборот.

На пути к фотонному компьютеру

Чтобы построить оптический компьютер, необходимо изготовить и соединить между собой большое количество различных элементов. В 70-х годах была разработана технология тонких пленок, прозрачных для лазерного излучения. Для взаимного соединения элементов оптических интегральных схем в принципе можно использовать волноводы, однако работа в этом направлении пока не проводилась.

Новую оптическую технологию можно было бы использовать для построения оптического компьютера с архитектурой, или логической организацией, такой же, как и в обычной ЭВМ. Основным достоинством вычислительной машины на оптических элементах с обычной архитектурой было бы ее высокое быстродействие. Логические схемы из оптических элементов могут выполнять до 1000 млрд. логических операций в секунду, в то время как максимальная скорость электронных переключающих устройств не превышает 1 млрд. операций в секунду.

Оптические переключатели к тому же обладают свойствами, позволяющими рационализировать конструкцию вычислительных машин. Если через транзистор пропустить несколько токов, они смешаются; в интерферометре оптические пучки не смешиваются, и поэтому один кристалл может выполнять несколько различных переключательных операций. К примеру, в кристалле одновременно может быть пять оптических каналов, образуемых разными лазерными пучками, и в каждом канале могут выполняться самостоятельные операции. Все операции при этом могут быть разными, если соответствующим образом подобраны интенсивности входных пучков. Например, два канала могут выполнять функцию И, два — функцию ИЛИ и один канал может работать в режиме обычного транзистора.

Если модулированные пучки во всех пяти каналах образовать из одного сигнала, разделив его на пять частей, то с этим сигналом можно выполнить одновременно пять различных логических операций. Полученные сигналы можно снова обработать либо последовательно, либо одновременно. Такой режим работы потребует совершенно иной организации машины, не похожей на организацию ЭВМ.

Применение оптических переключателей может привести и к иным изменениям конструкции и логики вычислительных машин. Как уже отмечалось, этому может способствовать появление устройств, имеющих больше двух устойчивых состояний. Некоторые кристаллы, с которыми мы проводили опыты, обладают способностью несколько раз скачкообразно повышать уровень пропускания при плавном увеличении интенсивности падающего пучка. Эти уровни могут определять состояния многозначных логических систем.

Оптическая обработка данных

Не следует думать, что оптические устройства просто придут на смену электронным устройствам. Наоборот, наибольшую пользу от них следует ожидать там, где их не смогут заменить никакие другие средства. Оптические волокна, обладающие способностью передавать колоссальные объемы информации, все шире используются для связи, включая и линии связи между вычислительными машинами. Если вычислительные операции будут выполняться оптическими устройствами, то, очевидно, и связь между вычислительными системами лучше всего осуществлять с помощью оптических волокон.

Мы уже говорили, что оптические переключатели можно изготавливать в виде тонких кристаллов. Современная техника выращивания кристаллов и тонкопленочная технология позволяют получать тонкие листы очень большой площади. Если оптическое изображение спроецировать непосредственно на такой большой лист, каждый элемент которого представляет собой оптический переключатель, то на выходе можно получить запись спроецированного изображения в цифровой форме. Получив такую запись, можно выполнять ее дальнейшую обработку различного вида, в том числе и увеличение изображения.

Мы в достаточной мере раскрыли

многообещающие потенциальные возможности и преимущества оптических переключающих устройств. Однако следует иметь в виду, что на пути к промышленной реализации результатов экспериментальных работ лежит немало трудностей. В своих исследованиях мы столкнулись с проблемой работы при низких температурах. Создание элементов, работающих при комнатной температуре, значительно упростило бы конструкцию и эксплуатацию оптических вычислительных машин. Выше говорилось, что ширина запрещенной зоны антимонида индия мала и уменьшается при повышении температуры. С помощью лазера на окиси углерода с длиной волны 10 мкм нам удалось получить оптическое переключение при комнатной температуре.

Без сомнения, новые лазеры и новые вещества с нелинейными оптическими свойствами позволят увеличить быстродействие и эффективность оптических переключателей. Вопрос создания материалов с необходимыми свойствами является ключевым в проблеме реализации оптической вычислительной машины. На практике желательно, чтобы переключатель потреблял как можно меньшую мощность. Энергоемкость оптических переключающих устройств измеряется произведением интенсивности падающего на кристалл света на площадь лицевой поверхности кристалла. Отсюда следует, что сокращение размеров устройства снижает потребляемую мощность. Однако устройство не может быть меньше площади сечения пучка лазера, которая ограничена возможностью фокусирующей системы. Кроме того, с уменьшением размеров устройства возникает проблема перегрева.

Следует отметить и другой факт — стремление к уменьшению мощности, затрачиваемой на переключение, противоречит условию повышения быстродействия, которое в свою очередь прямо зависит от энергии светового излучения на входе оптического ключа. При решении вопросов создания конструктивных элементов для оптических вычислительных машин размеры переключающих устройств, скорость срабатывания и потребляемая мощность должны рассматриваться с учетом их взаимного влияния и связи с оптическими и тепловыми свойствами нелинейной среды и функциями переключателя. Выбор подходящего вещества и поиски новых веществ с нелинейными характеристиками являются, пожалуй, самой главной задачей в создании оптических переключателей. И все же, несмотря на значительные трудности, мы считаем, что создание оптического компьютера —перспектива вполне реальная и многообещающая.

50 и 100 лет назад

SCIENTIFIC AMERICAN

ФЕВРАЛЬ 1933. Всегда считалось, что звезды, кажущиеся покрасневшими, находятся на очень большом расстоянии от Земли. Однако на состоявшемся в Мичиганском университете заседании Национальной академии наук д-р Дж. Стеббинс из обсерватории Маунт-Вилсон заявил, что некоторые красные звезды, находящиеся в полосе Млечного Пути, в четыре раза ближе к нам, чем считалось раньше. Слой «космической» пыли, который делит Млечный Путь на две полосы, приводит к покраснению звезд, находящихся за этим слоем. Точно так же Солнце кажется красным на закате, когда его лучи попадают к нам, пройдя более длинный путь в запыленной атмосфере нашей планеты.

Возможно, придет время, когда ракета станет очень опасным военным оружием. Такого мнения придерживаются многие авторитетные специалисты. Мощное устройство, управляемое по радио и способное разрушать объекты на большом расстоянии, гораздо быстрее начнет использоваться в военных целях, чем для аэронавигации. Быть может, столь большой интерес, проявляемый немецкими учеными к ракетостроению, объясняется тем, что Германии запрещено строить военные самолеты. Во всяком случае, по размаху исследований в этой области Германия опережает все другие страны. Недавно на берлинском аэродроме Темпельхоф была успешно продемонстрирована новейшая модель ракеты, сконструированная молодым инженером Тиллингом из Оснабрюка. Ракета имеет длину около 9 футов и сделана целиком из алюминия, а ее корпус завершается четырьмя хвостовыми стабилизаторами. В нижней части корпуса, там где крепятся стабилизаторы, внутри размещен двигатель, работающий на специальной медленно сгорающей пороховой смеси. Чтобы ракета совершила посадку в пределах аэродрома, она была заправлена небольшим количеством топлива — всего 12 фунтов. Пороховую смесь воспламенили электрической искрой с некоторого расстояния, и реактивная сила, возникающая за счет выброса выхлопных газов, подняла ракету на высоту около 2600 футов. При спуске с помощью простого гидравлического устройства из стабилизаторов ракеты постепенно автоматически выдвигались два крыла. Совершая парящее движение по спирали, ракета приземлилась всего в 400 ярдах от места старта.

SCIENTIFIC AMERICAN

ФЕВРАЛЬ 1883. На одном из последних заседаний Французской академии медицинских наук г-н Були от имени г-на Пастера сделал сообщение о выводах, к которым пришел выдающийся ученый, исследуя заболевание бешенством. Во-первых, было отмечено, что и паралитическая форма бешенства, и буйное бешенство, короче, все формы этого заболевания вызываются одним и тем же вирусом. Во-вторых, Були сообщил о таком опыте: в 1881 г. трем собакам был введен возбудитель бешенства, две собаки вскоре погибли, третья же, у которой проявились только признаки начальной стадии заболевания, выжила; в 1882 г. ей дважды вводили возбудитель заболевания под твердую мозговую оболочку, но собака не заболела. Сейчас в лаборатории Пастера имеются четыре собаки, которые невосприимчивы к заболеванию бешенством, каким бы способом им не вводили заведомо вирулентный материал. Пастер считает, что эти факты — первые шаги на пути к открытию эффективного метода предупреждения бешенства с помощью прививок. Однако он признает, что до окончательного решения проблемы еще далеко.

Недавно во французских газетах появилось сообщение о том, что полиция города Бордо запретила демонстрацию одного из аттракционов выставки — манекена-автомата Аз Ра. Было обнаружено, что этот манекен приводится в действие не автоматически, а человеком. Юноша 18 лет, здоровье которого было серьезно подорвано каждодневным непосильным трудом, скрывался в ящике автомата позади его механической части. Этот автомат напоминает знаменитого «турка-шахматиста», сконструированного в Венгрии в 1769 г. бароном Кемпеленом. Изобретение демонстрировалось в Германии, России, Франции, Англии и Америке, причем никому так и не удалось разгадать секрет действия автомата. В 1819-1820 гг. некто Мельцер вновь демонстрировал его в Англии. В 1844 г. Робер Гудин видел его в доме конструктора Кронуара из Бель-Виля. С тех пор судьба этого экспоната неизвестна, но вполне вероятно, что автомат Аз Ра есть не что иное, как «турок-шахматист».

Скрытые зрительные процессы

Зрение обычно считают единым чувством. Эксперименты показывают, однако, что зрительная система состоит из нескольких подсистем, деятельность которых в нормальных условиях скрыта от сознания наблюдателя

ДЖЕРЕМИ М. ВУЛФ

С ТОЧКИ зрения человека, размышляющего над своей способностью воспринимать окружающий мир, зрение — единое чувство. Распределение освещенности на изображениях, попадающих на сетчатки двух глаз, дает начало единому осознанию внешних объектов: их величины, формы, цвета, текстуры и ориентации в пространстве. Это мнение, однако, ошибочно. Зрительная система (т.е. механизм мозга, который перерабатывает информацию от глаз) в действительности представляет собой совокупность специализированных подсистем, в каждой из которых более или менее независимо от остальных подсистем обрабатывается определенная часть зрительной информации. Более того, результат действия некоторых подсистем зрительной системы трудно выявить. Они вносят свой вклад в работу мозга и в наше восприятие окружающего мира, но никакое самонаблюдение не может заставить нас осознать, почувствовать эти подсистемы как таковые. В них осуществляются скрытые зрительные процессы.

Как можно выявить скрытые процессы? Один путь — исследовать способности людей с поврежденным мозгом. Рассмотрим глазной зрачок, который сокращается в ответ на возрастание интенсивности света, падающего на сетчатку. Если у человека повреждена зрительная кора (так называют ту часть коры мозга, которая первой получает информацию от сетчаток глаз), он становится перцептивно слепым, т.е. чувствует себя незрячим так же, как и человек, потерявший возможность пользоваться собственно глазами. Тем не менее его зрачки продолжают сокращаться в ответ на действие света. Еще более поразительное явление обнаружили Э. Поппел, Р. Хелд и Д. Фрост из Массачусетского технологического института. Когда людей, перцептивно слепых из-за повреждений коры мозга, просят направить глаза на пятно света, они прекрасно это делают. Испытуемые говорят, что они не видят пятна и просто угадывают направление. Но они смотрят в правильном направлении чаще, чем можно объяснить случайностью.

Эти наблюдения указывают на множественность зрительных процессов: какие-то подсистемы повреждены, а другие, скрытые, дееспособны. Однако такие исследования не могут подтвердить, что и в отсутствие повреждений у людей с нормальной зрительной системой действуют те же скрытые процессы. Для этого требуются специальные экспериментальные приемы. Я опишу три серии экспериментов, каждая из которых выявляет зрительный процесс, в норме скрытый и недоступный самонаблюдению. В каком-то смысле все мы похожи на пациента с поврежденным мозгом, который может посмотреть на пятно света, но не воспринимает его. Мы тоже не осознаем существования некоторых звеньев зрительной системы, хотя наше поведение часто базируется на результатах их деятельности.

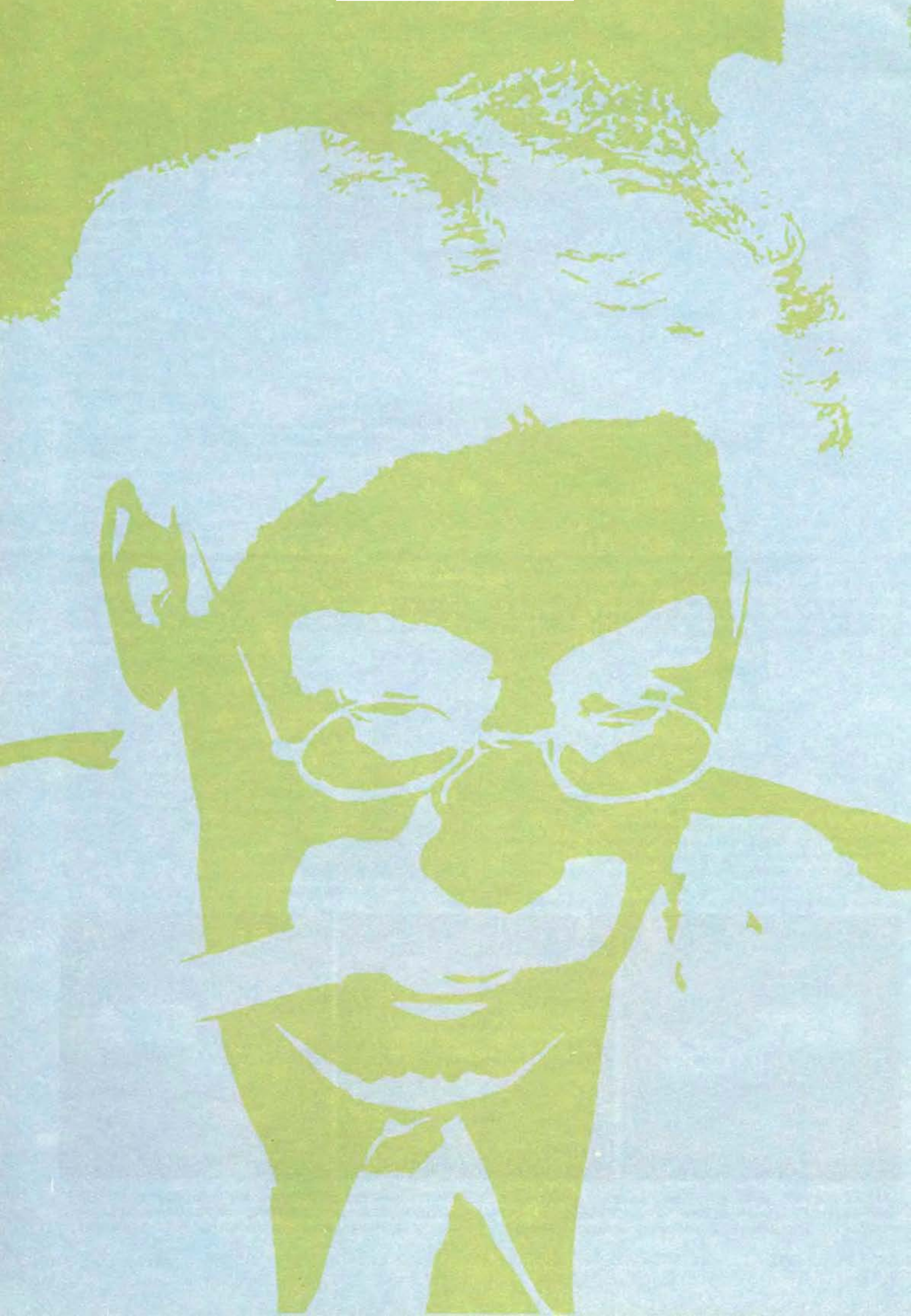
ОДНА из функций зрительной системы — управление мышцами, которые фокусируют глаз на различно удаленные объекты, изменяя форму хрусталика. Чем ближе объект, тем больше должна быть кривизна поверхности хрусталика. Изменение формы хрусталика называется зрительной аккомодацией. Судя по всему, мы способны аккомодировать на все, что можем увидеть. Напрашивается вывод, что аккомодации и зрительному восприятию доступна одна и та же группа стимулов.

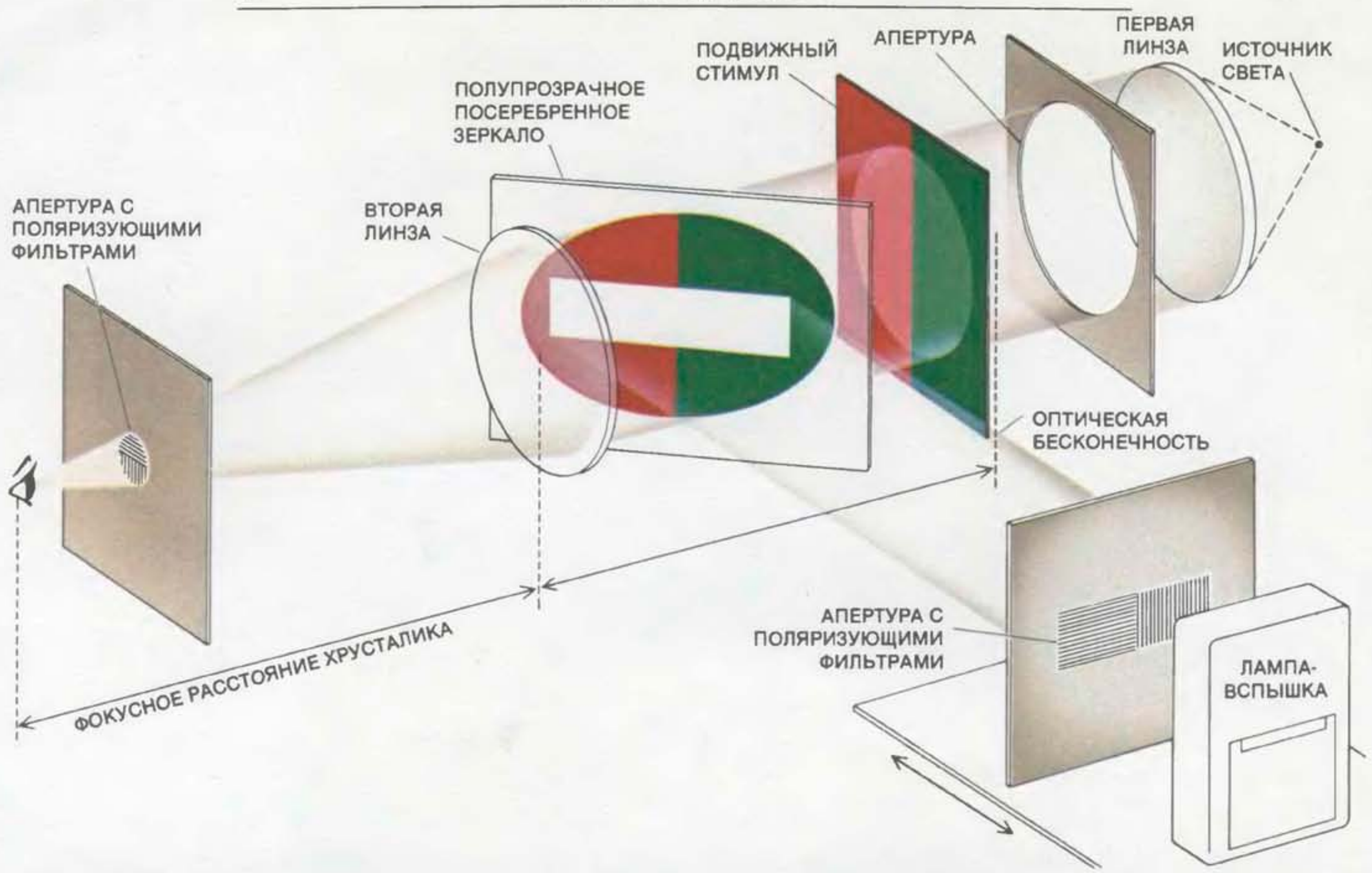
Так ли это? Есть несколько экспериментальных методик, с помощью которых можно оценить аккомодационное состояние глаза в то время, когда наблюдатель смотрит на тестовое изображение, удаленное на некоторое расстояние. В соответствии с одной такой методикой наблюдатель должен смотреть через поляризующие фильтры, а вспышка света из-за щели на короткое время создает на тестовом изображении яркую горизонтальную полосу. Фильтры не влияют на тестовое изображение, но заставляют свет от левой половины полосы проходить только через верхнюю часть хрусталика глаза наблюдателя, а свет от правой половины — через нижнюю.

Предположим, наблюдатель смотрит на тестовое изображение и аккомодирует на расстояние, большее или меньшее, чем расстояние до щели. Если при этом полоса вспыхивает, на сетчатке две ее половины окажутся не на одной прямой и наблюдатель увидит ступеньку. Теперь экспериментатору нужно, чтобы наблюдатель смотрел на стимул, а когда будет появляться полоска, отвечал, на одной ли прямой находятся ее половинки. Щель отодвигают на различные расстояния до тех пор, пока наблюдатель не сообщит, что половинки полосы встали на одну прямую. Расстояние от щели до глаза считается тем расстоянием, на которое аккомодирует наблюдатель.

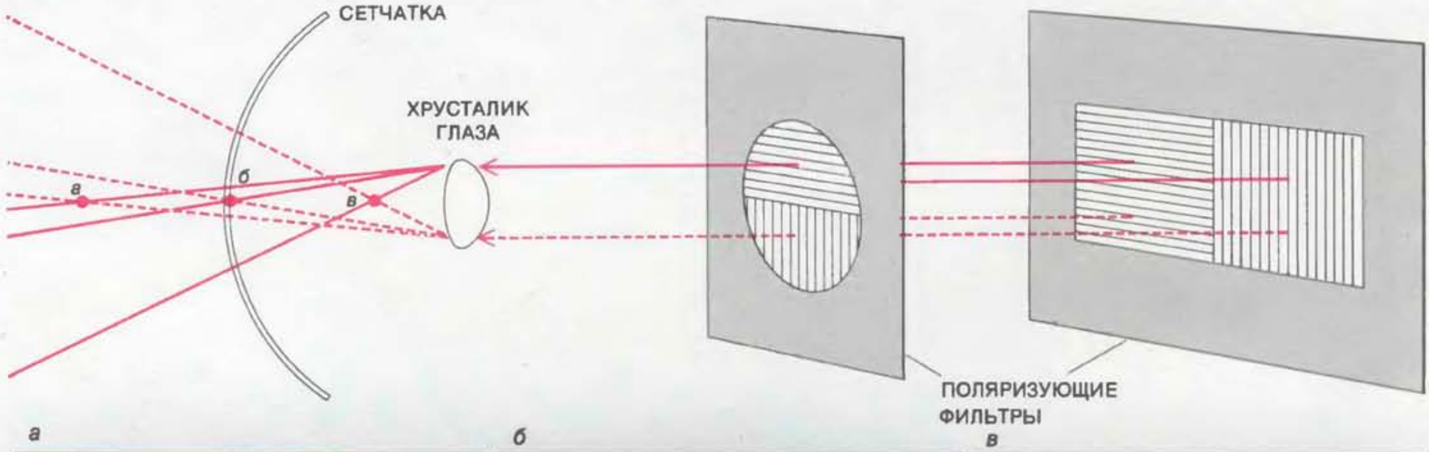
Хрусталик глаза твердеет с возрастом, и поэтому лучше изучать аккомодацию у людей не старше сорока лет. В лаборатории можно менять только расстояние до стимула, а его яркость и размер оставлять постоянным. При этом мы исключаем все воспринимаемые признаки расстояния до стимула, за исключением самого факта, что есть определенное состояние аккомодации и стимул сфокусирован. В таких условиях типичная аккомодация наблюдателя составляет около 90% от наилучшей (при наилучшей аккомодации сти-

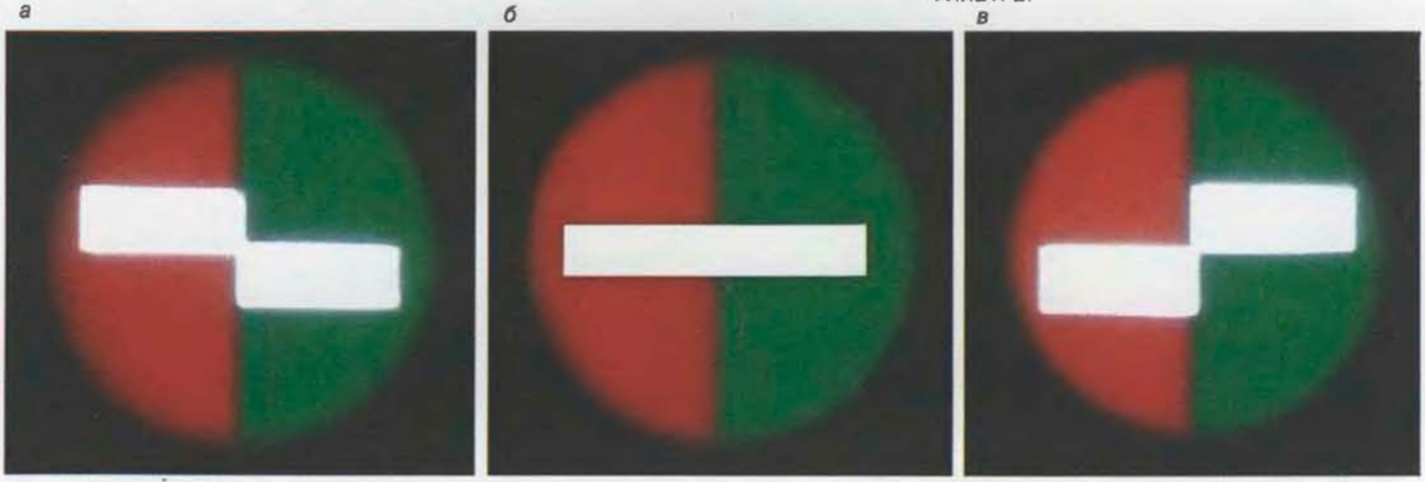
РАВНОМЕРНЫЙ ПО ЯРКОСТИ СТИМУЛ — это изображение, края в котором определяются только разницей в цвете. Здесь стимул далек от совершенства: голубые и зеленые участки изображения одинаковы по яркости лишь в той мере, в какой это позволяет типографская печать. Кроме того, яркость меняется на краю страницы, к тому же читатель держит журнал на расстоянии, удобном для чтения. В условиях лаборатории эти "указатели" можно устранить; тогда наблюдатель, глядя на стимул, теперь равномерный по яркости, окажется не в состоянии сфокусировать края стимула. Именно вследствие такого эффекта порой трудно узнать знакомое лицо. Эксперимент показал, что механизмы мозга, обеспечивающие зрительную аккомодацию (фокусировку глаз), не могут "видеть" цвет; это и есть скрытые механизмы, отличные от процессов, ведущих к восприятию. На фотографии вы видите актера Граучо Маркса в фильме "Лошадиные перья".





ЭКСПЕРИМЕНТАЛЬНАЯ УСТАНОВКА для проверки способности глаза аккомодировать на край в стимуле, равномерном по яркости. Первая линза собирает свет в пучок, он проходит через апертуру и стимул. Вторая линза направляет свет в глаз наблюдателя. На короткое время на стимуле вспыхивает горизонтальная полоска света (ее создает приспособление из лампы-вспышки, подвижной щели и полупрозрачного посеребренного зеркала).





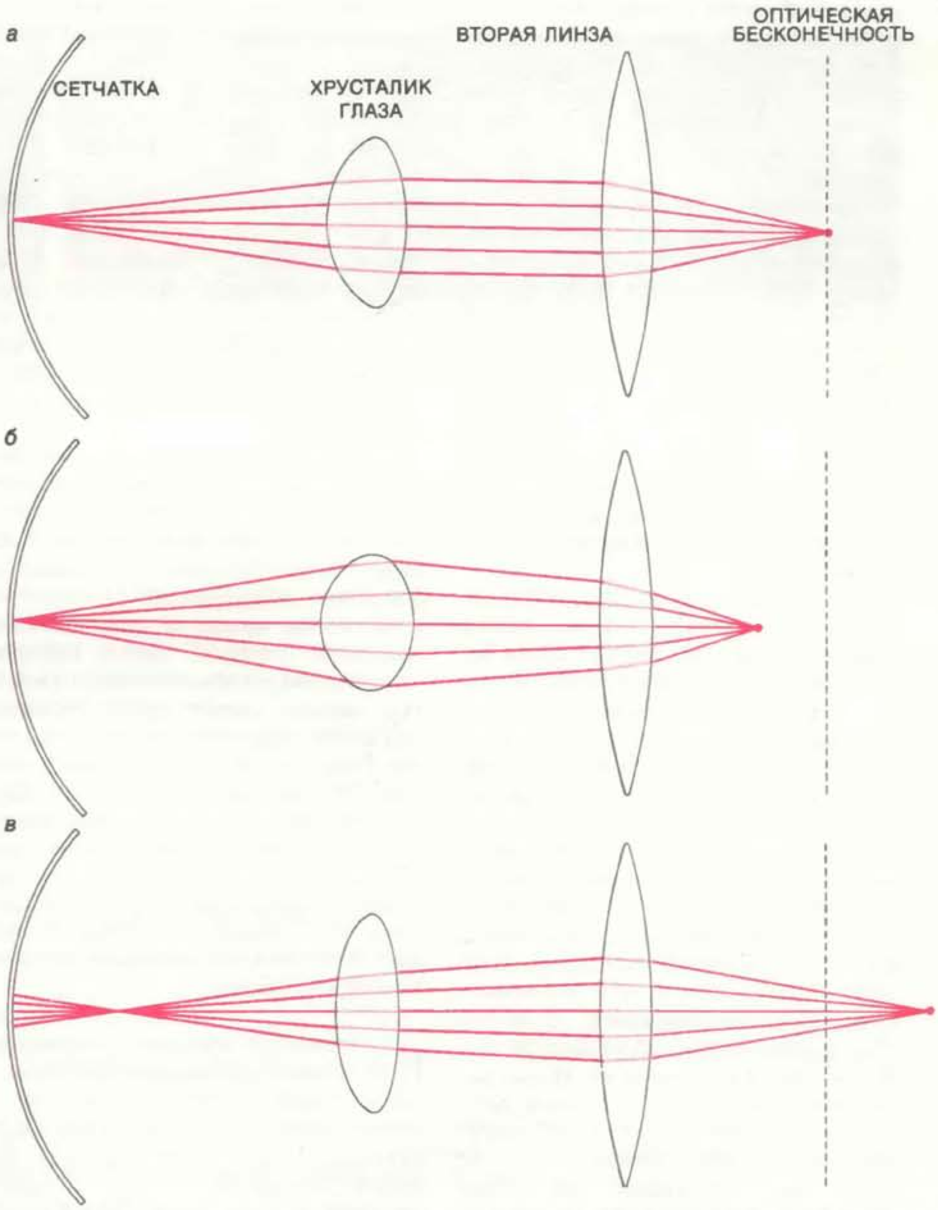
АККОМОДАЦИЮ измеряют, предъявляя наблюдателю вспыхивающую горизонтальную полоску. Благодаря серии поляризующих фильтров свет от левой ее половины проходит только через верх хрусталика глаза, свет от правой — только через нижнюю его часть. Человек увидит ровную линию, если аккомодирует на расстояние подвижной щели.

мул, удаленный на 25 см, вызвал бы аккомодацию глаза на расстояние 25 см).

Как реагирует глаз, если у стимула нет признаков, на которые глаз мог бы сфокусировать? Например, если вас окружает белый экран, ощущение такое, будто вы находитесь внутри гигантского шарика для пинг-понга. Другой пример — белый экран меньшего размера, если на него смотреть через линзу, которая заставляет экран казаться слишком близким (по оптическому, или кажущемуся, расстоянию), так что глаз не может сфокусировать его края. Сюда же можно отнести чистое небо и полную темноту. Г. Лейбович и Д. Оуэнс из Пенсильванского университета обнаружили, что в этих условиях хрусталик глаза принимает устойчивую кривизну покоя, соответствующую специфическому фокусному расстоянию, которое называется темновым фокусом. У каждого человека свой характерный темновой фокус, обычно его величина около 1 м, что примерно равно расстоянию от ладони вытянутой руки до глаза.

То же самое происходит, когда стимулом служит решетка из белых и черных линий, таких тонких, что они воспринимаются как серое поле. С помощью линзы экспериментатор опять-таки может располагать стимул на различных оптических расстояниях. Когда стимул находится на оптическом расстоянии 25 см, хрусталик глаза наблюдателя приходит в состояние покоя. Когда стимул отодвигают на оптическое расстояние 1 м, состояние аккомодации наблюдателя не изменяется. Следовательно, при аккомодации форма хрусталика глаза изменяется, чтобы держать объект в фокусе, только если системе, ответственной за аккомодацию, предъявляют что-то, что она может "увидеть".

ЗНАЯ УЖЕ, что система аккомодации слепа к некоторым аспектам окружающего мира, Д. А. Оуэнс и я задались вопросом, может ли эта система "видеть" цвет. Мы измеряли аккомодацию у наблюдателя, глядевшего на простой стимул — круглое поле, разделенное по вертикали пополам так, что был виден один вертикальный край. Благодаря оптической системе, называемой вид Максвелла, периметр круга находился за фокусным расстоянием ближайшей к глазу линзы. Это значит, что лучи света от любой точки на периметре круга входили в глаз таким образом, что их невозможно было свести в фокус, какую бы форму ни принял хрусталик глаза наблюдателя. В результате, кроме вертикального края между двумя половинами поля, в тестовом изображении не было никакого другого края, на который глаз мог бы аккомодировать. Оптическую систему приспособили так, чтобы размер изображения на сетчатке оставался одним и тем же, когда оптическое расстояние стимула меняли.



ОПТИЧЕСКОЕ РАССТОЯНИЕ до стимула — это видимое расстояние, придаваемое стимулу второй линзой экспериментальной установки. Если стимул расположен на фокусном расстоянии от линзы (*а*), он находится относительно глаза в оптической бесконечности, т.е. лучи света от всех частей стимула поступают в глаз параллельно, как если бы они исходили от бесконечно удаленного источника. Кривизна хрусталика глаза станет такой, что лучи сойдутся в одну точку на сетчатке. Если стимул расположить ближе к линзе, чем ее фокусное расстояние, поступающие в глаз лучи будут расходящимися (*б*). Чтобы лучи сошлись в фокус, форма хрусталика должна еще более приблизиться к сферической. Если стимул расположить дальше фокусного расстояния, лучи будут конвергировать на пути к глазу (*в*). Кривизна хрусталика глаза не сможет уменьшиться в нужной степени, и лучи сойдутся не достигнув сетчатки, т.е. стимул нельзя привести в фокус.

Край можно было бы создать, сделав так, чтобы между двумя полуполями была разница в цвете либо в яркости. В реальном мире большинство контуров возникает благодаря различиям по обоим параметрам. Мы сделали одну половину поля красной, а другую зеленой и варьировали яркость одной или другой половины. Таким образом мы получили набор стимулов: наша шкала начиналась с края между красным и черным и кончалась краем между черным и зеленым. В двух крайних случаях контраст по яркости между красным (или зеленым) и черным был 100%. В середине шкалы красное и зеленое поля были одинаково яркими; в таком равномерном по яркости стимуле край создавался только разницей в цвете, а контраст по яркости был нулевым.

Когда мы проверяли способность людей аккомодировать на каждый из этих стимулов, мы обнаружили, что она уменьшается, когда контраст по яркости слабеет. В наших опытах с равномерным по яркости стимулом наблюдатели отмечали, что хорошо

видят край между красным и зеленым, но не могут его сфокусировать. В одном эксперименте мы предъявляли наблюдателям черную букву *E* на белом фоне. Наблюдатели аккомодировали достаточно хорошо на оптическом расстоянии от бесконечности до 22 см. Затем мы предлагали им смотреть на край между красным и черным и на край между черным и зеленым. Аккомодация составила 80% от наилучшей, достигнутой в опыте с буквой *E*. В заключение мы показывали им край между красным и зеленым без контраста по яркости. Аккомодация составила только 19% от наилучшей. Когда мы повторили эксперимент с такими парами цветов, как красный и оранжевый или голубой и зеленый, показатели наблюдателей были не лучше или еще хуже.

Как и любой другой оптической системе, глазу свойственна хроматическая аберрация: свет разного цвета фокусируется в несколько различных точках. Вследствие аберрации на сетчатке изображение полуполя одного цвета сдвинуто относительно изображения другого полуполя. В результате может возникнуть светлый или темный контур (в зависимости от того, разделены стимулы или перекрываются). Поэтому очень трудно сделать стимул, в котором впечатление края создавалось бы исключительно за счет различия по цвету. Если бы нам удалось исключить хроматическую аберрацию, в наших опытах показатели наблюдателей были бы, вероятно, еще хуже. Итак, вывод кажется ясным: аккомодация цветнослепа или по крайней мере "видит" цвет очень плохо. Аккомодация как подсистема существенно обособлена от зрительного восприятия. В отношении механизмов зрения самонаблюдение, стало быть, бесполезно.

ДРУГОЙ подход при изучении подсистем зрительной системы состоит в том, чтобы рассматривать способы, при помощи которых данная зрительная функция выполняется различными частями системы. Одна из таких функций — бинокулярное зрение. В норме у человека два глаза, и мозг объединяет информацию от них. Сейчас уже ясно, однако, что мозг не просто суммирует ви́дение двумя глазами в одном бинокулярном процессе. Есть, видимо, несколько специализированных механизмов, которые комбинируют информацию от двух глаз каждый своим, особенным образом. Так же, как никаким самонаблюдением не обнаружить, что аккомодация цветнослепа, так нельзя осознать, что бинокулярное восприятие состоит из многих процессов. Тем не менее эти процессы существуют и их можно выявить в эксперименте.

Рассмотрим зрительную иллюзию, которая в той или иной форме знакома каждому. Допустим, вы сидите в поезде, ожидая, когда он тронется. В окно вы видите другой, неподвижный поезд. Вот этот другой поезд поехал назад, а вам кажется, будто вперед движется ваш поезд. Только через какое-то время вы понимаете, что отправился соседний поезд, а ваш все еще стоит на месте. Движение другого поезда каким-то образом ввело вас в заблуждение. Иллюзия движения возникла исключительно благодаря зрительному раздражителю. В таких случаях говорят об индуцированном движении, т.е. наведенном ощущении движения.

Некоторые исследователи пытались изучить механизмы мозга, ответственные за иллюзию движения. Хелда и меня интересовал несколько иной вопрос. Мы хотели узнать, что́ механизм ощущения движения может "видеть", в частности, нужно ли для ощущения движения смотреть двумя глазами. Казалось бы, чтобы получить ответ на этот вопрос, следует просто сравнить ощущение движения, когда открыты оба глаза и когда открыт только один глаз. Обычно, однако, для возникновения максимального ощущения движения достаточно смотреть одним глазом. Значит, нужен такой стимул, у которого бинокулярный эффект больше, чем монокулярный, т.е. в стимуле должно быть нечто недоступное восприятию одним глазом.

КОМПОНЕНТ, который удовлетворяет этому требованию, называется циклопическим стимулом — в честь гомеровского одноглазого Циклопа, к которому попал в плен Одиссей. Название подходящее, потому что циклопический стимул виден, только если мозг объединяет информацию от двух глаз, т.е. действует как один глаз Циклопа. Циклопическим стимулом, например, является небольшое различие в положении изображений объекта на сетчатках разных глаз. Выявить его очень просто. Вытяните руку и посмотрите на кончик пальца сначала одним глазом, а потом — другим. Вы сразу заметите разницу. Благодаря таким различиям мозг воспринимает глубину. Ясно, что мозгу нужны оба глаза, чтобы использовать циклопический стимул.

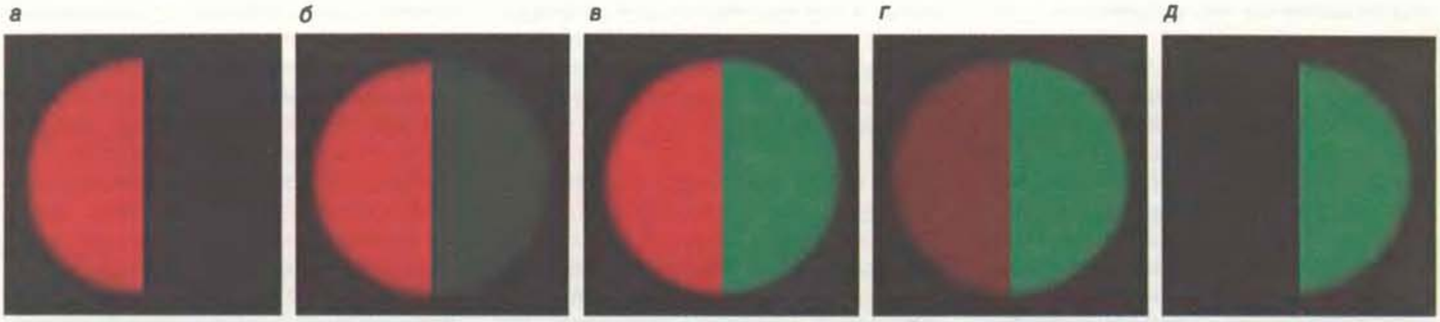
В наших опытах циклопический стимул для наведения ощущения движения основан на хорошо известном факте, что на киноэкране объекты кажутся плавно двигающимися, хотя нам показывают последовательность неподвижных фотографий. Испытуемый сидел внутри цилиндра диаметром 1 м и высотой 1,7 м. На белой внутренней поверхности цилиндра были беспорядочно разбросаны черные кружочки диаметром 2,5 см. Цилиндр вращался вокруг испытуемого со скоростью 30 градусов в секунду. Когда внутреннюю поверхность цилиндра освещали обыкновенной лампой, испытуемому казалось, что он вращается в противоположном направлении. Мы пользовались стробоскопической лампой — такая лампа освещает объект периодическими вспышками; при каждой вспышке перед глазами испытуемого возникал, так сказать, кадр из фильма о черных точках. При этом человек ощущал, что движется.

Для того чтобы создать циклопический стимул, понадобились две стробоскопические лампы. Одну закрыли красным фильтром, а другую — зеленым. Испытуемый надевал разноцветные очки, в которых перед одним глазом помещался красный фильтр, а перед другим — зеленый. Через зеленый фильтр не мог пройти красный свет, а через красный — зеленый свет. Поэтому свет от одной стробоскопической лампы был виден только одному глазу, свет от другой — только другому.

Допустим, каждая стробоскопическая лампа вспыхивает с частотой 10 Гц. Если лампы вспыхивают синхронно, испытуемый не должен заметить никакой существенной разницы между рассматриванием двумя глазами и рассматриванием одним глазом. В обоих случаях он увидит "фильм" о движущихся точках, скорость проекции которого 10 кадров в секунду. Допустим теперь, что стробоскопические лампы работают в противофазе, т.е. вспыхивают попеременно. Открыв только один глаз, испытуемый снова увидит "фильм" со скоростью проекции 10 кадров в секунду. Открыв оба глаза, он увидит другой "фильм" — со скоростью проекции 20 кадров в секунду, который состоит из кадров, предъявленных попеременно сначала одному глазу, потом — второму, затем снова первому и т.д.

Этот "межглазной фильм" и есть циклопический стимул: его нельзя увидеть одним или другим глазом по отдельности. Он создает впечатление движения стимула. Вопрос в том, "видит" ли его система ощущения собственного движения. Если да, то в опыте со стробоскопическими лампами в противофазе должно возникать более сильное ощущение собственного движения, чем в опыте с лампами, работающими в фазе. Если это не так, в обоих случаях впечатление должно быть одинаковым.

СОБРАВ нужную аппаратуру, мы стали готовить к экспериментам испытуемых. В частности, мы обучали их методу, известному в экспериментальной психологии как оценка величины. Мы просили испытуемых давать оценку 10 при максимальном ощущении собственного движения и оценку 0



НАБОР СТИМУЛОВ для исследования способности к аккомодации: от края между красным и черным (*а*) до края между черным и зеленым (*д*). Стимул *в* равномерен по яркости. Круглый край каждого стимула определяется апертурой, расположенной далеко за фокусным расстоянием второй линзы оптической установки, поэтому глаз наблюдателя не может привести этот край в фокус.



СПОСОБНОСТЬ ФОКУСИРОВАТЬ стимул зависит от контраста по яркости. Здесь приведены средние показатели четырех испытуемых (в процентах от способности, которую они продемонстрировали при рассматривании черных букв на проверочной таблице). Показатели тем хуже, чем меньше контраст, и минимальны (19% от лучшего результата), когда стимул равномерен по яркости.

при отсутствии ощущения движения. Оценки между 0 и 10 нужно было давать промежуточным ощущениям. За 25 лет исследователи в области экспериментальной психологии (особенно отметим С. Стивенса из Гарвардского университета) убедились, что, как ни странно, люди очень верно оценивают свои ощущения: результаты воспроизводятся и вполне надежны. Нас беспокоило, что разницы в результатах двух экспериментов может не быть, если ощущение собственного движения будет оцениваться максимально, то есть в 10 единиц, когда лампы работают в фазе. Если частота кадров "фильма" постепенно уменьшается, иллюзия движения изображения становится все более слабой. Для ощущения собственного движения важна частота, с которой вспыхивают лампы. Мы обнаружили, что при любой частоте мелькания больше 2,5 Гц возникает более или менее сильное ощущение движения, но только при частотах больше 15 Гц испытуемые давали своему ощущению оценку 10. Значит, частоты от 2,5 до 15 Гц годились для наших целей. Циклопический стимул при вспыхивании ламп в противофазе способен, в принципе, создать более сильное ощущение движения, чем стимул, который создают лампы, вспыхивающие синхронно (частота мелькания вдвое больше). Результаты наших опытов оказались однозначными. При частоте мелькания от 2,5 до 15 Гц испытуемые давали всегда бóльшую оценку, когда лампы зажигались и гасли в противофазе. Очевидно, система ощущения собственного движения может "видеть" чисто бинокулярный стимул. Означает ли этот результат, что в мозгу есть несколько бинокулярных механизмов? Видимо, нет. Как я отмечал, небольшая разница в положении изображений объекта на сетчатках двух глаз представляет собой циклопический "ключ" к восприятию глубины. Можно предполагать, что бинокулярный механизм, обслуживающий этот аспект зрительного восприятия, участвует также в возникновении ощущения движения.

Некоторые люди, однако, не воспринимают глубину по циклопическим стимулам. Например, они не испытывают иллюзии глубины, когда глядят в стереоскоп или смотрят стереофильмы. Они стереослепы. У одних это врожденное, подобно тому как можно от рождения не различать цвета. У других во младенчестве развились такие дефекты, из-за которых зрительные оси при фиксации далеких объектов непараллельны. Стереослепота встречается реже, чем цветовая слепота: в США от нее страдают всего несколько процентов населения. Стереослепой человек может все же воспринимать глубину, исходя из видимой величины знакомых объектов или из того, что одни объекты в поле зрения заслоняют другие, расположенные дальше. Эти монокулярные признаки вполне достаточны для ощущения глубины, и многие стереослепые люди даже не подозревают о своем дефекте, пока

это не выявится при специальной проверке.

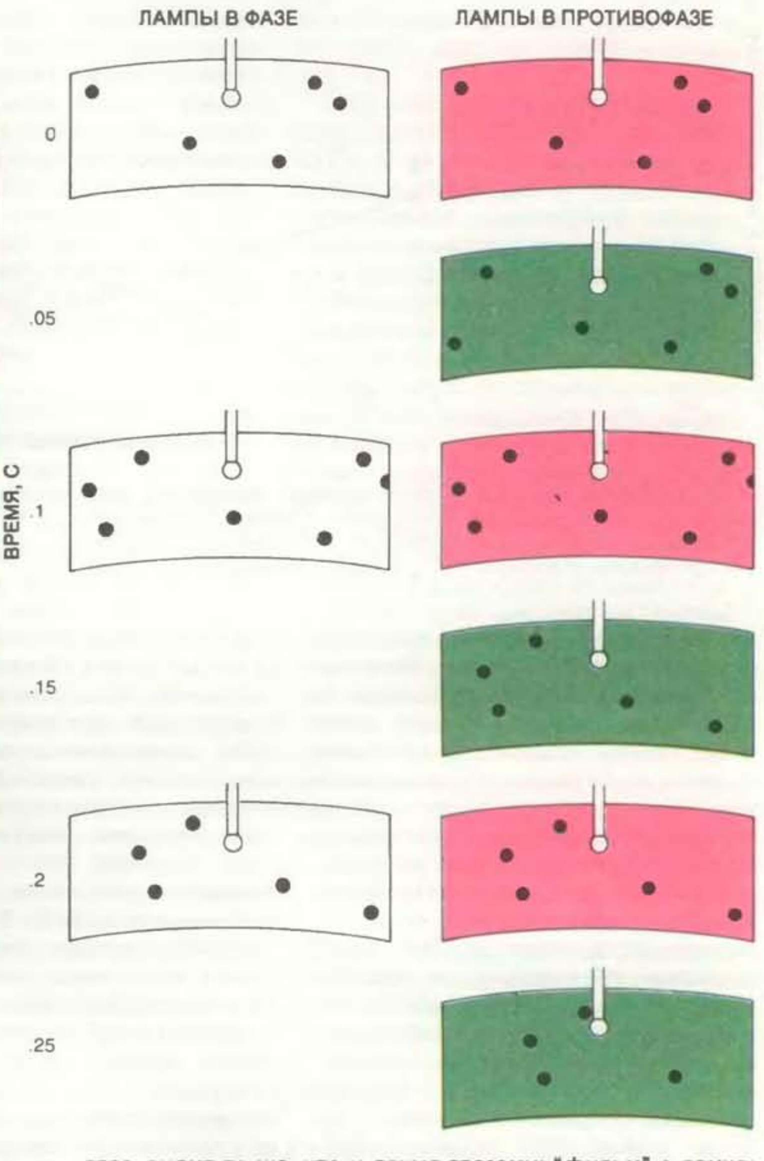
Хелд и я попросили четырех человек, у которых выявили стереослепоту, принять участие в наших экспериментах по ощущению движения. Они оценивали свои ощущения так же, как и нормальные испытуемые. У них тоже ощущение собственного движения было больше при наличии чисто бинокулярного стимула. Как видим, и в этом случае люди с поврежденным бинокулярным механизмом в экспериментах, требовавших использования именно бинокулярного механизма мозга, вели себя совершенно нормально. Следовательно, механизм, поврежденный у стереослепых, не единственный бинокулярный механизм в мозгу. Должны быть по крайней мере два бинокулярных механизма: один обеспечивает стереоскопическое восприятие глубины, а другой участвует в возникновении ощущения движения. Снова эксперименты показали, что зрительная информация перерабатывается более чем в одной системе.

ДВА СКРЫТЫХ процесса, которые я сейчас описал (цветнослепой зрительный процесс, от которого зависит аккомодация, и бинокулярный зрительный процесс, участвующий в возникновении ощущения движения), по-видимому, не имеют прямого отношения к зрительному восприятию. Другие скрытые процессы принимают участие именно в восприятии.

Для примера я вернусь к бинокулярным зрительным процессам. Наши эксперименты по исследованию ощущения движения выявили два бинокулярных процесса. Эксперименты, о которых пойдет речь, показали большее. В самом деле, выявляется большой набор механизмов. Между "глядением" и единым восприятием лежат процессы, которые "видят" только одним глазом — либо левым, либо правым, процессы, которые "видят" любым глазом, и, что наиболее удивительно, чисто бинокулярные процессы, которые "видят" только двумя глазами одновременно, причем в последнем случае оба глаза должны смотреть на один и тот же стимул. Если бы в зрительной системе был только чисто бинокулярный механизм, мы бы не могли видеть, закрыв один глаз. Например, вы бы ничего не увидели в такой ситуации: вы смотрите на дом из-за дерева так, что один глаз видит только ствол, а другой глядит вдаль.

Способность воспринимать чисто бинокулярный стимул (такой, как циклопический стимул, который мы ис-





ИЛЛЮЗИЯ СОБСТВЕННОГО ДВИЖЕНИЯ возникает, когда наблюдатель сидит внутри вращающегося цилиндра, на белой внутренней поверхности которого нарисованы в беспорядке черные точки. Две стробоскопические лампы освещают поверхность цилиндра вспышками красного и зеленого цвета; благодаря очкам свет каждого цвета поступает только в один глаз. Если лампы вспыхивают в фазе с частотой 10 вспышек в секунду, наблюдатель, у которого открыт один глаз, видит то же, что и двумя глазами: "фильм" о движущихся точках со скоростью проекции 10 кадров в секунду. Если лампы вспыхивают в противофазе, наблюдатель, открыв один глаз, видит тот же "фильм", а открыв два глаза — другой, со скоростью проекции 20 кадров в секунду. Этот "фильм" межглазной: его последовательные кадры воспринимаются разными глазами попеременно. Стационарное световое пятно служит точкой фиксации.

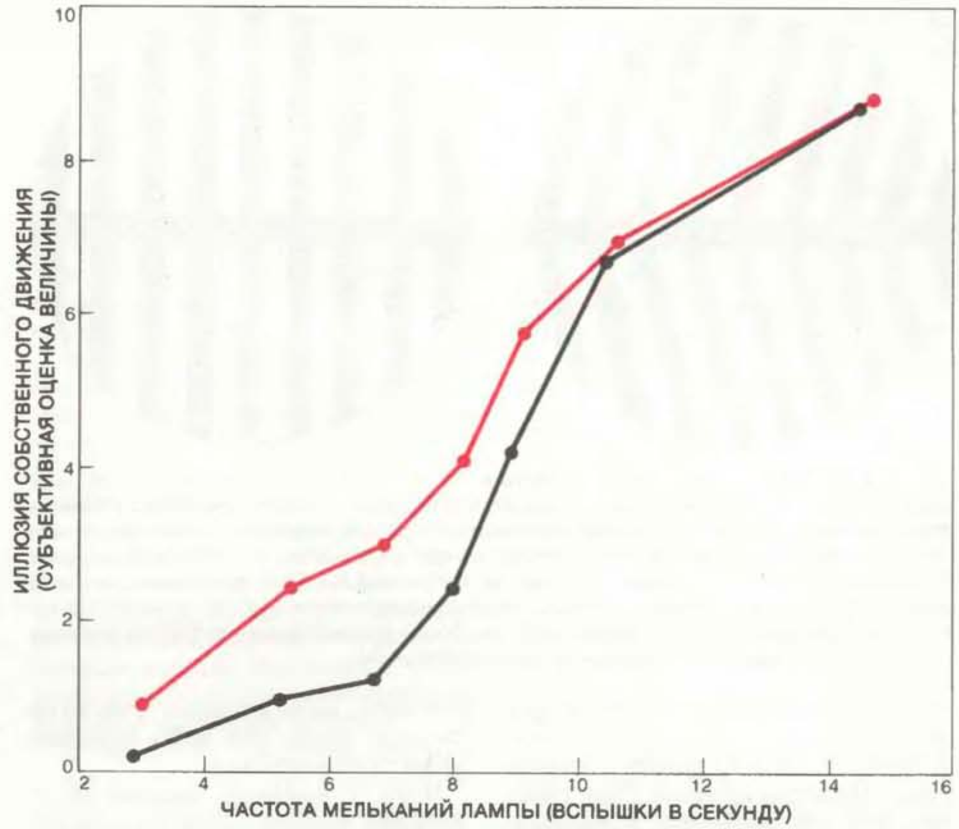
пользовали в экспериментах по исследованию ощущения движения), не означает, что в мозгу есть чисто бинокулярный зрительный механизм. Возможно, что циклопический стимул "видят" зрительные механизмы, которые принимают информацию как от обоих глаз, так и от каждого глаза в отдельности. Как же тогда показать, что чисто бинокулярные процессы действительно существуют? Снова прибегнем к косвенному методу. Известно временное искажение увиденного, которое называют последействием наклона.

Посмотрите на узор узких полосок справа на рисунке на с. 34. Верхняя половина узора должна казаться коллинеарной нижней половине. Теперь посмотрите на узор наклонных линий слева и одну-две минуты двигайте глазами взад-вперед вдоль горизонтальной черной полосы, которая пересекает центр левого узора. Затем быстро переведите взгляд на правый узор. Вам покажется, что его половины не коллинеарны, а наклонены в сторону, противоположную направлению наклона на левом узоре.

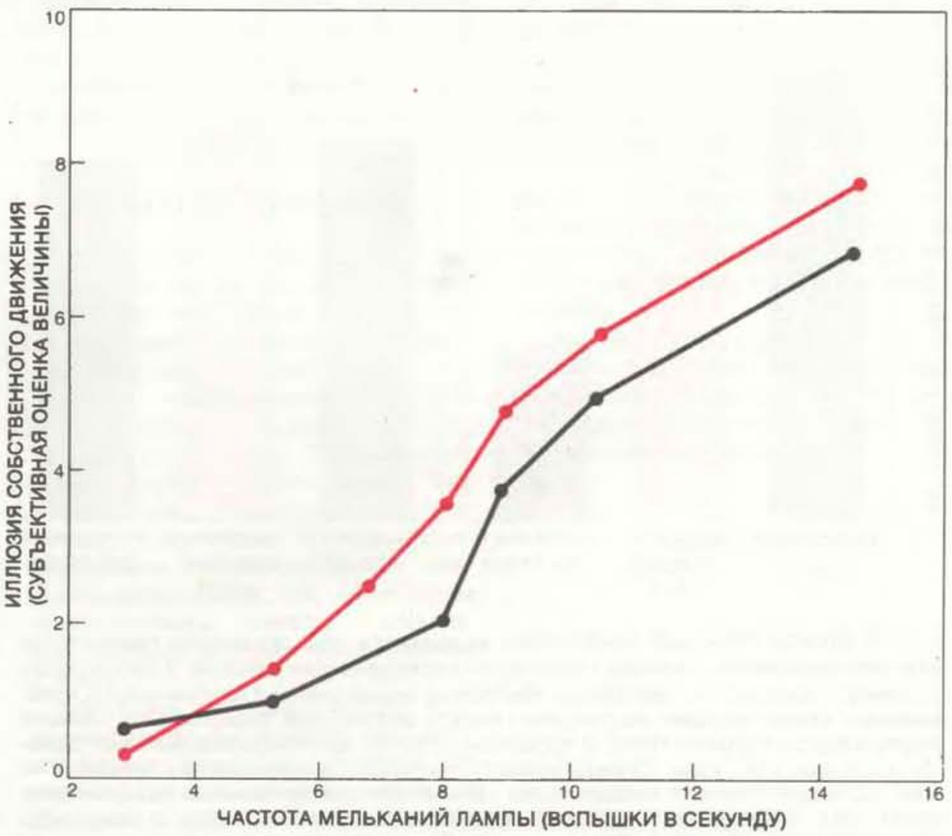
Это и есть последействие наклона. Его можно измерить: для этого половины правого узора должны быть подвижными — тогда наблюдатель сможет привести их в такое положение, чтобы они ему казались коллинеарными. До рассматривания левого узора испытуемые устанавливают полоски точно коллинеарно, а после — непременно ошибаются примерно на 2°.

ПРЕДПОЛОЖИМ, наблюдатель смотрит на узор наклонных линий только правым глазом, а затем подгоняет половины правого узора, глядя на него только левым глазом. Левый глаз не видел наклонных линий, тем не менее обнаруживается последействие наклона. Происходит так называемый межглазной перенос. Очевидно, исследуемый механизм бинокулярный, но не чисто бинокулярный. Он начинает действовать, когда правый глаз смотрит на наклонные линии, а также когда левый глаз проверяют подвижным узором. По-видимому, механизм может отвечать на раздражение как левого глаза, так и правого.

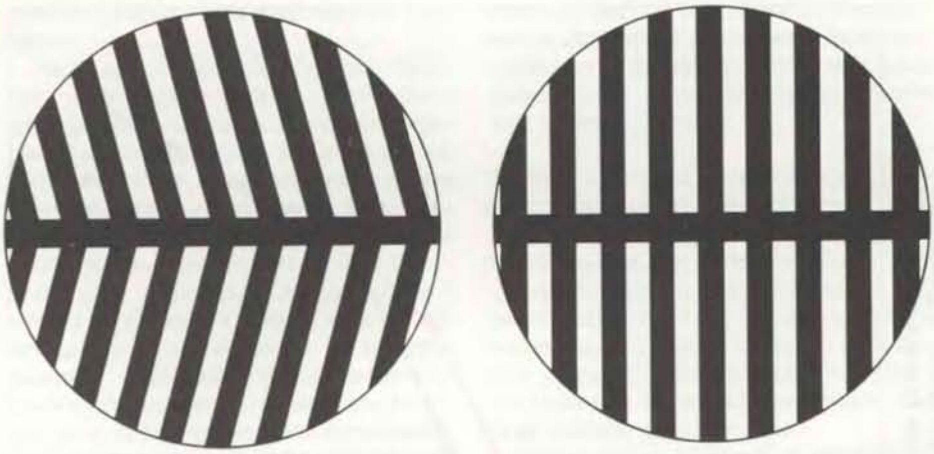
Оказалось, что последействие наклона выражено слабее, когда его проверяют на глазе, которому не показывали наклонные линии, чем когда его проверяют на глазе, видевшем этот узор. Эффект переносится только на 70 — 80%. Р. Блэйк и его сотрудники из Северо-Западного университета считают, что, если последействие наклона проверяют на одном глазе, действуют по крайней мере два механизма. Один из них бинокулярный, а другой — монокулярный. Когда наклонные линии



ОЩУЩЕНИЕ СОБСТВЕННОГО ДВИЖЕНИЯ сильнее, если лампы вспыхивают в противофазе (цветная кривая), чем при вспышках в фазе (черная кривая). Это свидетельствует о том, что механизм мозга, ответственный за ощущение движения, "видит фильм", возникающий при попеременной стимуляции глаз. Значит, этот механизм бинокулярный. В эксперименте участвовали 6 испытуемых.



ОЦЕНКА ВЕЛИЧИНЫ ощущения, проделанная стереослепыми испытуемыми, показала, что у них ощущение собственного движения сильнее, если лампы вспыхивают в противофазе. Следовательно, бинокулярный механизм, ответственный за ощущение движения, отличен от того, которого лишены стереослепые.



ПОСЛЕДЕЙСТВИЕ НАКЛОНА — иллюзия, которая дает возможность изучать скрытые зрительные процессы. Попробуйте поводить некоторое время глазами взад и вперед вдоль горизонтальной черной полосы, пересекающей наклонные линии на левом рисунке. После этого вам на миг покажется, что полосы в верхней и нижней половинах правого рисунка не коллинеарны. Последействие наклона измеряют, сделав половины правого узора подвижными и предлагая наблюдателю отрегулировать его так, чтобы сразу же после рассматривания левого рисунка линии на правом рисунке казались коллинеарными.

показывают правому глазу, начинают действовать оба механизма — бинокулярный и монокулярный правого глаза. Проверка правого глаза снова включает оба механизма. В результате — сильное последействие наклона. Предположим теперь, что проверяется левый глаз. Бинокулярный механизм "видел" стимул и включается. Монокулярный механизм левого глаза тоже действует, но он не "видел" узор из наклонных линий. Его вклад ослабляет эффект последействия.

Итак, с помощью явления последействия наклона Блэйк продемонстрировал, что имеют место два процесса — монокулярный и бинокулярный. Хелд и я использовали последействие, чтобы показать, что наряду с ними действует и чисто бинокулярный механизм. Мы предлагали испытуемым смотреть на наклонные линии двумя глазами, но попеременно: одну минуту левым глазом, одну минуту правым глазом, затем снова одну минуту левым и одну минуту правым. Измеряя последействие, их просили смотреть либо левым глазом, либо правым, либо обоими глазами одновременно.

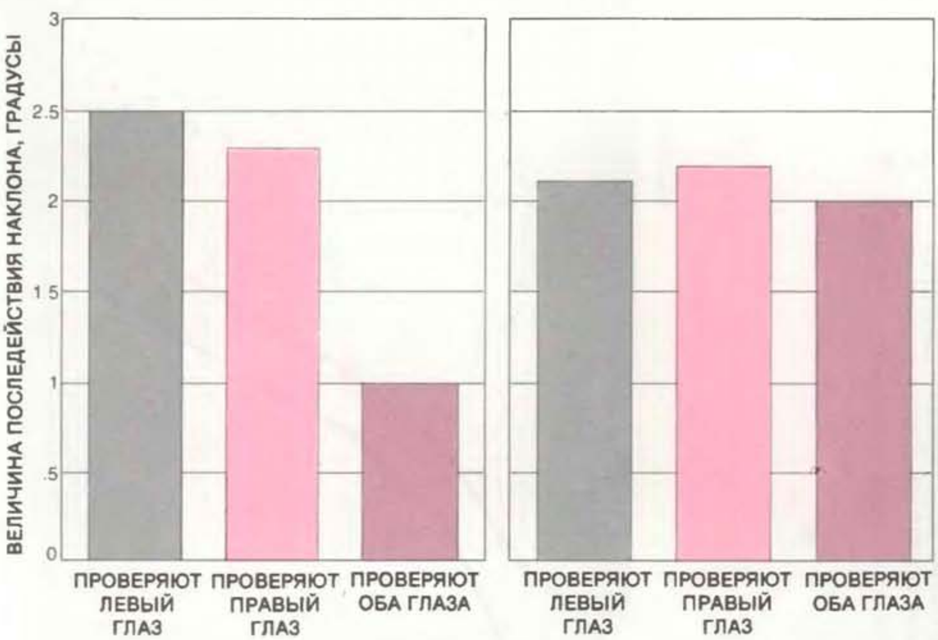
Мы полагали, что попеременное монокулярное рассматривание стимула выявит и тот зрительный процесс, который использует информацию от левого глаза, и тот, который связан с правым глазом, и тот, в котором учитывается информация от каждого глаза. Если бы это были единственные зрительные процессы, на них бы и строилась вся зрительная система. Тогда не должно быть разницы, открыт ли у наблюдателя один глаз или оба, когда он подгоняет половинки подвижного узора: последействие наклона будет одинаковым. Если же есть зрительные механизмы, действующие, только когда открыты оба глаза одновременно, картина будет иной. Попеременное монокулярное рассматривание стимула не сможет выявить эти процессы, но, если потом при проверке последействия открыть оба глаза, включится чисто бинокулярный процесс и эффект последействия наклона будет слабее.

Именно это мы и обнаружили. Эффект последействия наклона был гораздо меньше, когда испытуемый открывал оба глаза, чем когда он смотрел одним глазом. В другом опыте мы предлагали испытуемым смотреть на наклонные линии обоими глазами. Мы ожидали, что это выявит весь зрительный процесс, и действительно, последействие было теперь практически одинаковым, смотрел ли наблюдатель обоими глазами или только одним, когда подгонял половинки подвижного узора.

В обоих экспериментах и левому, и правому глазу показывали наклонные линии. Единственное различие состояло в том, что в первом опыте испытуемые не смотрели на них двумя глазами одновременно, а во втором смотрели. В первом эксперименте один зрительный процесс не выявлялся, но проявлялся во втором. Этот процесс "видит" только двумя глазами, то есть представляет собой чисто бинокулярный процесс.

Итак, мы можем видеть не только посредством чисто бинокулярного процесса, но и благодаря либо монокулярному процессу, либо процессу, "видящему" любым глазом.

Зрение следует рассматривать не как единое чувство, а как совокупность механизмов, каждый из которых можно считать чувством. Весь набор зрительных чувств человека нам еще неизвестен, впереди — новые открытия.



ЧИСТО БИНОКУЛЯРНЫЙ МЕХАНИЗМ выявляется при сравнении результатов двух экспериментов, в которых проверяли последействие наклона. В первом опыте (*слева*) наблюдатель смотрел на наклонные линии одним и другим глазом попеременно; таким образом, могли действовать зрительный процесс, получающий информацию от одного глаза, и процессы, получающие информацию то от одного, то от другого глаза. Стимул недоступен чисто бинокулярному процессу — тому, который получает информацию только при одновременном раздражении обоих глаз. Наблюдатель подгонял половины подвижного узора, открыв либо один глаз, либо оба глаза. Оказалось, что бинокулярное последействие меньше монокулярного. Второй эксперимент (*справа*) отличался только тем, что наблюдатель рассматривал наклонные линии двумя глазами одновременно; при этом бинокулярное последействие оказалось таким же, как и монокулярное, следовательно, выявился чисто бинокулярный процесс.

А было ли поднятие?

СООБЩЕНИЯ о подвижках земной коры вызывают живейший интерес среди жителей южной Калифорнии, где земля буквально иссечена разрывами. Особенно внимательны к подобного рода информации жители городка Палмдейл — небольшого аграрного центра, который расположен непосредственно в зоне разлома Сан-Андреас, в 55 км к северо-востоку от Лос-Анджелеса. В 1976 г. трое геологов писали в журнале "Science", что по сравнению с 1955 г. Палмдейл приподнялся почти на 30 см. Позднее один из авторов этого сообщения — Р. Касл из Геологической службы США — пришел к выводу, что величина поднятия на самом деле была даже больше 30 см. Кроме того, Р. Касл отметил, что половина поднятия пришлась, по-видимому, на период с 1961 по 1965 г. Эти подвижки нельзя было объяснить каким-либо местным сейсмическим явлением, однако не исключалась возможность того, что они свидетельствуют о близящемся землетрясении.

С тех пор постоянно предлагались различные гипотезы о причинах появления палмдейлского «выступа», однако ни одна из них не выглядела убедительной. Но вот недавно С. Холдал из Национального управления по исследованию океанов и атмосферы высказался в том духе, что никакого объяснения и не требуется, поскольку в действительности сколько-нибудь значительного поднятия вовсе не было. Пытаясь согласовать свою точку зрения с противоречащими ей результатами других исследователей, С. Холдал ссылается на практический опыт съемщиков и цитирует своего коллегу, геофизика У. Стрейнджа, который в 1981 г. сказал: «В юго-западных районах США, для которых характерен высокий уровень солнечной радиации, не следует производить нивелирование вдоль железнодорожных путей».

В статье, опубликованной в журнале "Journal of Geophysical Research", С. Холдал описывает, как могло возникнуть ошибочное представление о поднятии близ Палмдейла. Прибор, которым до 1964 г. пользовались в Калифорнии для измерения высот на местности, был удобен для ведения наблюдений на больших расстояниях. Обычно съемщики производили измерения вдоль железнодорожного полотна, где уклон земной поверхности невелик, что и позволяло визировать на значительных расстояниях. Удобно было и то, что с одной точки на другую съемщики переезжали на дрезине. В то же время еще в 30-е годы знали, что результаты нивелирной съемки могут искажаться из-за рефракции: градиент температуры воздуха вдоль линии визирования вызывает искривление световых лучей, вследствие чего замеры превышений оказываются неточными. Тем не менее в Соединенных Штатах поправки на рефракцию в данные нивелирных съемок не вносились, поскольку считалось, что величина ошибки незначительна.

Вероятно, в большинстве случаев ошибка действительно невелика. Однако и материал железнодорожных насыпей, и сами рельсы сильно поглощают тепло, так что поблизости от них устанавливаются существенные градиенты температуры воздуха; это особенно касается солнечных юго-западных районов США. Таким образом, ошибка вследствие рефракции при нивелировании вдоль железнодорожного полотна была велика, а протяженность интервалов визирования еще более осложняла ситуацию.

В 1964—65 гг. появились нивелиры новой конструкции, что позволило сократить интервалы визирования и, следовательно, уменьшить ошибку из-за рефракции даже в тех случаях, когда съемка производилась в полосе отчуждения вдоль железной дороги. С. Холдал внес поправку на рефракцию в те данные нивелирных съемок, на основании которых говорили о поднятии земной поверхности в районе Палмдейла. Амплитуда поднятия оказалась не 30 см, а менее 7,5 ± 4 см. Но даже и эту величину можно отнести за счет остаточной ошибки в измерениях.

Восприятие топологии

СОВРЕМЕННЫЕ представления о механизме зрения исходят из утверждения, что мозг анализирует изображение, выделяя в нем некие абстрактные признаки. Замечательные тому примеры — обнаружение края и обнаружение топологических свойств, таких, как связанность. Согласно сообщению журнала "Science", Лин Чен, работающий в Китайском университете науки и техники в Хофее и в Калифорнийском университете в Сан-Диего, предполагает, что обнаружение топологических свойств — древнейшая основная функция зрительной системы.

Основания для такого вывода дали эксперименты, показавшие, что человек легко выделяет фигуру из фона, даже если детали стимула неясны и аморфны, а значит, мозг при этом реагирует только на те свойства стимула, которые имеют общий характер, т.е. присущи ему как целому. Такие свойства, отмечает Чен, можно описать математически — как топологические свойства, например связанность. Топологически круг, треугольник и квадрат эквивалентны. Любые две точки такой фигуры можно соединить отрезком прямой, который целиком поместится внутри фигуры. В этом смысле круг, треугольник и квадрат принципиально отличаются от любой фигуры с отверстием, например от кольца.

В одном из экспериментов испытуемые нажимали на ключ, для того чтобы увидеть стимул, состоящий из круга и квадрата, круга и треугольника или круга и кольца. Каждый стимул появлялся всего на 5 миллисекунд; затем наблюдателя спрашивали — одинаковые или разные фигуры ему показывали. Оказалось, что при такой постановке опыта зрительная система лучше всего различает круг и кольцо: испытуемые назвали их разными в 64,5% случаев. Круг и квадрат были «разными» в 43,5% случаев, а круг и треугольник — только в 38,5%.

В другом эксперименте испытуемым предъявляли короткий отрезок прямой, который освещался двумя вспышками света. Между первой и второй вспышками положение и наклон прямой изменяли, и наблюдатель должен был сказать, какой отрезок ему показали первым, а какой — вторым. Испытуемые отвечали правильно в 59,2% случаев. Тот же опыт ставили еще так: в стимул включали неподвижный круг, расположенный так, чтобы один из отрезков находился внутри него, а другой — вне. Казалось бы, неподвижный круг не давал никакой добавочной информации о порядке предъявления отрезков. Тем не менее испытуемые определяли последовательность предъявления отрезков правильно в 79% случаев. Очевидно, топологические признаки (в данном случае расположение отрезка внутри или вне замкнутой кривой) помогали идентификации.

«Мы не знаем, — замечает Чен, — как именно возникает реакция на топологические свойства объекта. Это принципиальная проблема в изучении зрительного восприятия».

Решеточная теория удержания кварков

С помощью дискретной решетки, налагаемой на пространство-время, удалось смоделировать силу, действующую между кварками в таких частицах, как протон, и объяснить, почему изолированный кварк не может быть в свободном состоянии

КЛАУДИО РЕББИ

РАЗВИТИЕ квантовой механики заставило критически отнестись к трактовке элементарных частиц как неких «кирпичиков», из которых «построено» вещество. Во многих случаях такие частицы вовсе не ведут себя как твердые, непроницаемые «кирпичики», а могут быть представлены в виде волн. Тем не менее до недавнего времени все еще казалось, что элементарные частицы подобны «кирпичикам», по крайней мере в том смысле, что в принципе их можно изолировать и наблюдать как самостоятельные объекты. Например, электрон, протон и нейтрон можно рассматривать отдельно как независимые волновые пакеты. Однако даже столь вольное толкование понятия «кирпичики» неприменимо, когда речь идет о кварках, из которых, как полагают, состоят протон, нейтрон и многие родственные им частицы. По-видимому, кварк не может быть изолирован. Несмотря на то что получены многочисленные доказательства существования кварков и антикварков, связанных в пáры и тройки, отдельный, или свободный, кварк никогда не наблюдался.

По мере накопления экспериментальных данных возникло представление, согласно которому даже если кварки и являются реальными частицами, то они всегда должны находиться в связанном состоянии внутри ядерных частиц. Любая теория взаимодействия кварков должна объяснять это явление, которое принято называть пленением или удержанием кварков *. Легко построить наглядную модель таких частиц, как протон, которые составлены из связанных между собой кварков. Можно представить, например, что кварки «привязаны» к концам струны, которую невозможно разорвать. Тогда они могут свободно двигаться только в некотором объеме, определяемом длиной струны, но не смогут находиться слишком далеко друг от друга. Однако построение теории, которая объясняла бы структуру ядерных частиц и тот факт, что кварки в них всегда «заперты» (находятся в связанном состоянии), без нарушения ограничений, налагаемых теорией относительности, квантовой механикой и обычным принципом причинности, — задача, грандиозная по своей сложности.

Сейчас, после нескольких лет экспериментов и теоретических исследований, большинство физиков убеждены, что у них наконец есть теория, способная объяснить взаимодействие кварков. Одна из причин такой уверенности кроется в том, что в математическом смысле она аналогична наиболее удачной физической теории из когда-либо созданных. Имеется в виду квантовая теория взаимодействий в электромагнитном поле — квантовая электродинамика (КЭД). Идейная общность КЭД и теории взаимодействия кварков отражена в самом названии последней: квантовая хромодинамика (КХД).

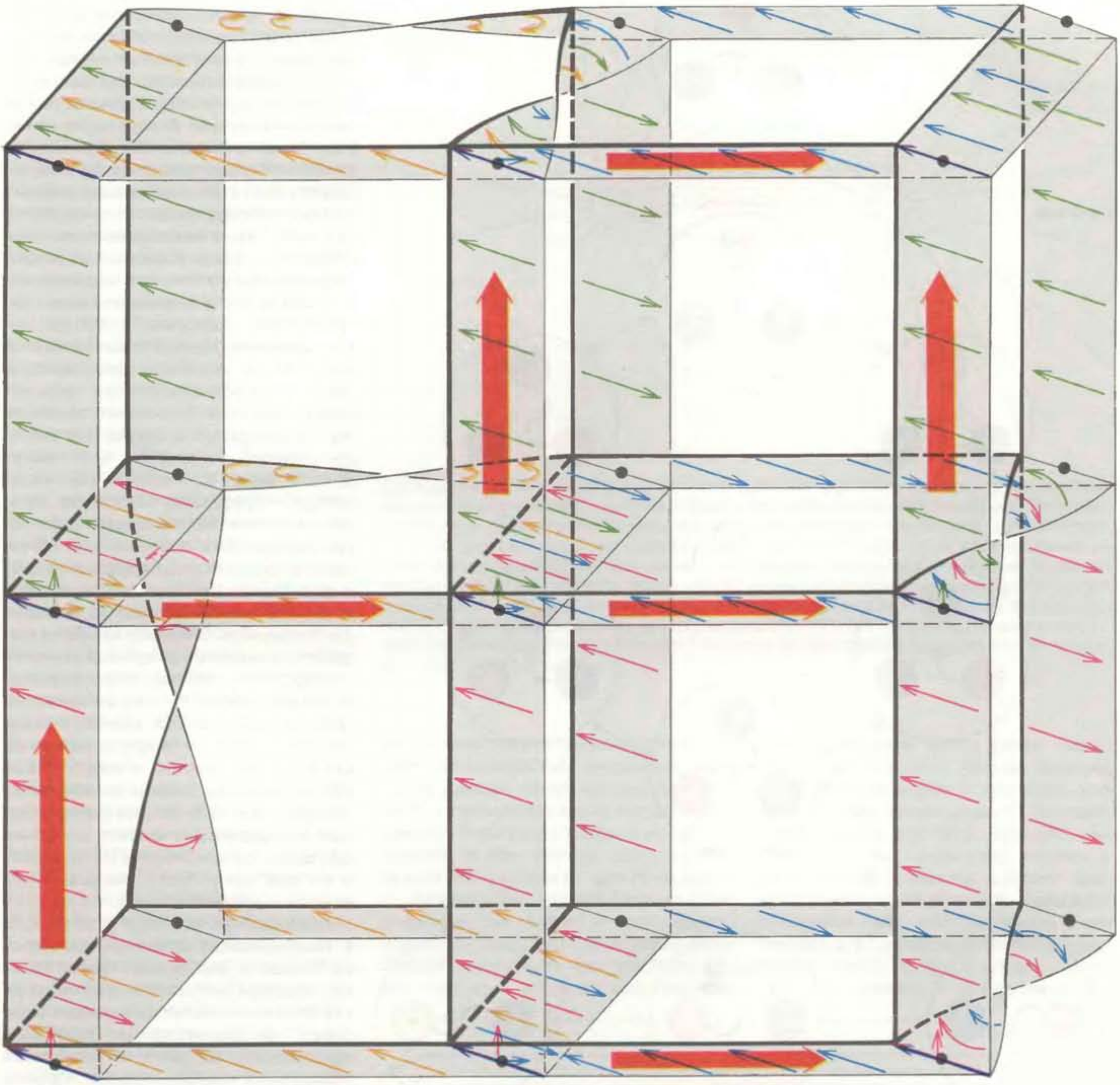
Окончательному признанию КХД препятствует сложность математических расчетов, которые затрудняют возможность точных предсказаний, получаемых аналитически на ее основе. Действительно, в рамках КХД до сих пор не доказано явление, которое физики давно хотят обосновать, — удержание кварков в элементарных частицах. Однако недавно мы с коллегами по Брукхейвенской национальной лаборатории применили для решения этой проблемы математические методы, опирающиеся на возможности ЭВМ. В результате был достигнут неожиданный успех. Этими методами свойства КХД исследуются с помощью последовательности все более точных приближений, поэтому результаты вычислений не имеют такой же силы, как логические выводы из основных принципов теории. Однако численные расчеты дают веские доказательства в пользу удержания кварков.

Структура вычислений основана на предложении, сделанном впервые в 1974 г. К. Вильсоном из Корнеллского университета. Вильсон предложил сформулировать КХД на кубической решетке — упорядоченной структуре, которая разбивает пространство-время на дискретные точки. Решетка является лишь приближением к реальному пространству-времени, но она позволяет делать расчеты, которые иначе выполнить невозможно. По мере того как ячейка решетки делается все меньше, значения физических величин, вычисленных на решетке, стремятся к значениям, которые могли бы быть получены на основе КХД в обычном, непрерывном, пространстве-времени. Наши приближенные численные расчеты показали, что для решетки с бесконечно малой ячейкой удержание кварков есть следствие КХД. По причинам, которые станут ясны ниже, и КХД, и КЭД называют калибровочными теориями, поэтому вычислительный метод, который я буду описывать, можно назвать калибровочной теорией на решетке.

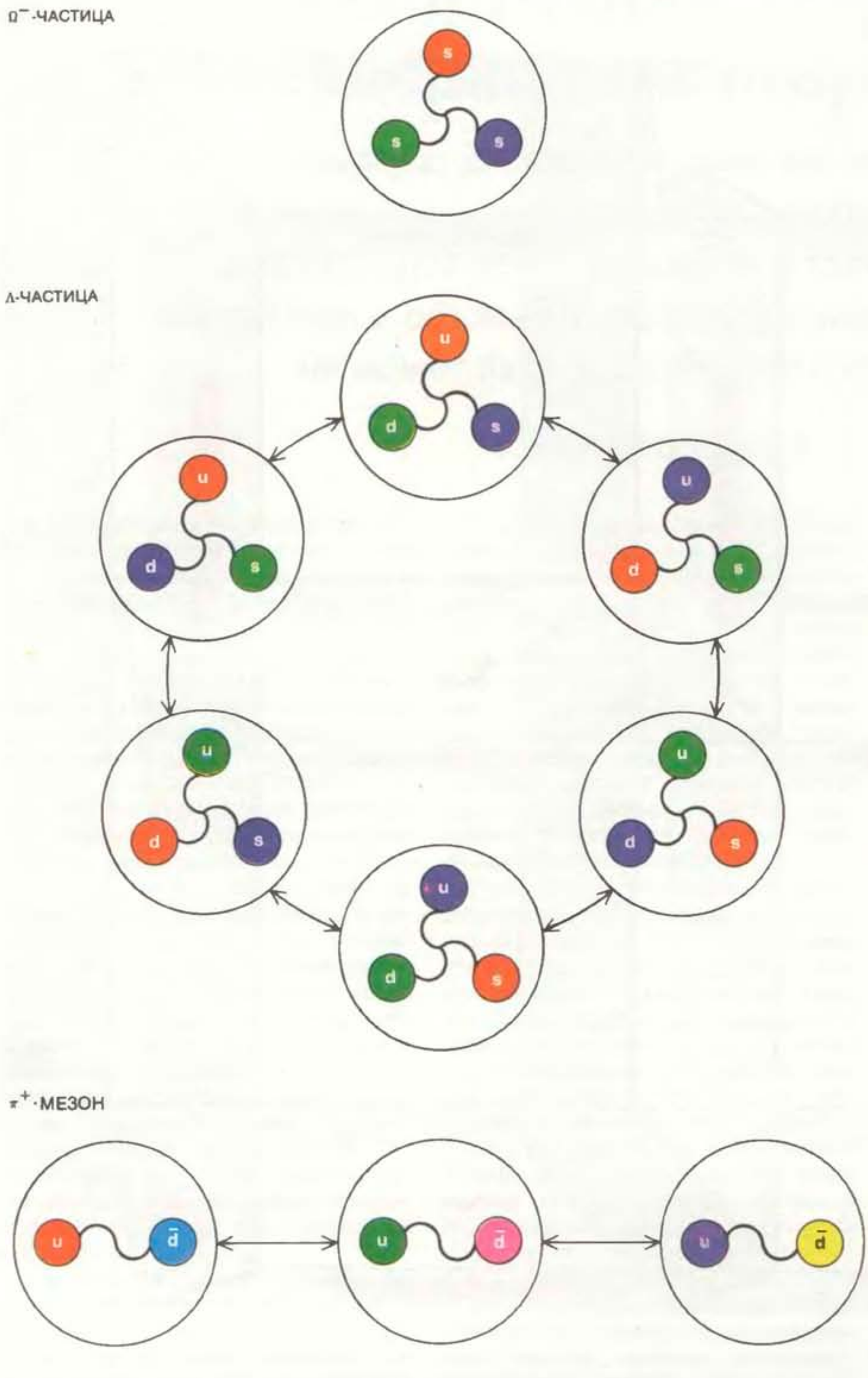
ТОЛЧКОМ к созданию кварковой модели послужила необходимость в систематизации большого числа частиц, которые участвуют в сильных взаимодействиях. Протон и нейтрон относятся к частицам этого типа: именно вследствие сильных взаимодействий они связаны в ядрах атомов. Заключение о существовании множества других сильно взаимодействующих частиц было сделано на основе анализа продуктов распада частиц, сталкиваемых в ускорителях. Время жизни большинства таких частиц чрезвычайно мало — порядка 10^{-24} с. Все частицы, участвующие в сильных взаимодействиях, называют адронами (от греческого *αδροζ* — крупный, массивный).

В 1962 г. М. Гелл-Манн из Калифор-

* В современной научной литературе не существует однозначного перевода английского термина «confinement»; иногда для описания удержания кварков в адронах используют транслитерацию «конфайнмент». — *Прим. перев.*



РЕШЕТКА, СОСТАВЛЕННАЯ ИЗ ТОЧЕК, — схематическое представление пространства-времени; с помощью такой решетки физики пытаются понять природу силового поля, постоянно удерживающего кварки в адронах. Кварки и антикварки, из которых, как полагают, состоят протоны, нейтроны и все остальные частицы, участвующие в сильных взаимодействиях, никогда не наблюдались в свободном состоянии — они связаны вместе в пáры и тройки. Поле, ответственное за удержание кварков, называют хромоэлектрическим; математически оно классифицируется как калибровочное поле. Калибровка подобна линейке или стрелке, она используется для сравнения физических величин, заданных в различных узлах решетки. Длина линейки и направление стрелки могут изменяться при их перемещениях в пространстве-времени. Частицы, например кварки, расположены только в точках (узлах) решетки. На рисунке фиолетовыми стрелками изображены переменные величины, определяющие данное состояние частицы. В каждом узле фронтальной плоскости решетки они должны иметь одно из двух направлений. Чтобы сравнить ориентацию двух фиолетовых стрелок, надо переместить одну из них вплотную к другой. Для этого необходимо ввести набор правил, определяющих изменения ориентации стрелки при ее переносе; эти правила изображены в виде лент, соединяющих соседние узлы. Набор таких правил и есть калибровочное поле. Оно позволяет сместить стрелку вверх или направо от узла и перенести назад в исходный узел вокруг грани — квадрата, образованного четырьмя узлами. (Стрелки, которые переносятся, окрашены в зеленый, желтый, голубой, малиновый цвета.) Набор лент, который возвращает переносимую стрелку к ее первоначальной ориентации, представляет состояние калибровочного поля с наименьшей энергией. Однако набор лент, напоминающий лист Мёбиуса, не может быть раскручен: в этом случае переносимая стрелка вернется в исходный узел, имея противоположную ориентацию. Энергия калибровочного поля на решетке сосредоточена именно на таких гранях. Красными стрелками показаны силовые линии калибровочного поля.



ЦВЕТ — квантовомеханическая характеристика кварков, введенная для того, чтобы согласовать кварковую модель с принципом Паули. До введения цвета казалось, что кварковая модель предсказывает существование частиц, для которых принцип Паули нарушается. Предполагалось, например, что составляющие Ω^--частицу три *s*-кварка находятся в одинаковом квантовомеханическом состоянии. Противоречие устраняется, если предположить, что кварки могут принимать один из трех цветов: красный, фиолетовый или зеленый. Теперь кварки имеют разный цвет, поэтому три кварка могут находиться в разных квантовомеханических состояниях и принцип Паули соблюдается. Цветные кварки должны быть скомбинированы так, чтобы Ω^--частица была бесцветной или белой. Поскольку составная частица бесцветна, для нее равновероятны любые допустимые комбинации цветов. Например, для Λ-частицы, состоящей из трех кварков, возможны шесть бесцветных состояний, соответствующих шести разным перестановкам трех цветов: красного, фиолетового и зеленого. Положительный π^+-мезон, составленный из кварка и антикварка, может быть в любом из трех бесцветных состояний: комбинаций красного и голубого («антикрасного»), зеленого и малинового («антизеленого»), фиолетового и желтого («антифиолетового») цветов. Кварки связаны между собой струной, которую невозможно порвать. Модель изображает удержание кварков в адронах.

нийского технологического института и Ю. Нееман из Тель-Авивского университета предложили схему для классификации адронов, объединяющую частицы в симметричные семейства. Схема основывалась на математической теории групп и была названа «восьмеричным путем». Вскоре после этого Гелл-Манн и независимо Дж. Цвейг, также из Калифорнийского технологического института, предложили физическую интерпретацию «восьмеричного пути». Такую математическую классификацию можно объяснить на основе предположения, что все адроны построены из более фундаментальных составляющих, которые Гелл-Манн назвал кварками. Каждый известный в то время адрон мог быть представлен в виде некоторой комбинации трех основных кварков (и соответствующих им антикварков): *u*-кварка (от англ. up — вверх), *d*-кварка (от англ. down — вниз) и *s*-кварка (от англ. strange — странный). Например, протон является комбинацией двух *u*-кварков и одного *d*-кварка, в то время как нейтрон есть комбинация *u*-кварка и двух *d*-кварков. Положительно заряженный π^+-мезон состоит из *u*-кварка и антикварка $\bar{d}$. С тех пор как была выдвинута гипотеза о кварковой структуре адронов, открыты новые адроны. Возникла необходимость добавить по крайней мере еще два кварка: *c*-кварк (от англ. charm — очарованный) и *b*-кварк (от англ. bottom — низ)*. Тем не менее кварковая модель по-прежнему остается в высшей степени удачной схемой классификации частиц: в настоящее время известно более 100 адронов, и все они могут быть описаны с помощью кварковой модели.

Но, несмотря на успехи этой модели в классификации адронов, некоторые особенности, изначально приписываемые кваркам, мешали воспринимать их как физически реальные частицы. Наиболее фундаментальная проблема здесь связана с неудачными попытками обнаружить свободный кварк. Протон и нейтрон сильно связаны в атомном ядре, тем не менее их можно высвободить, сталкивая ядра достаточно высоких энергий. Поэтому любая теория, описывающая взаимодействие кварков, должна не только объяснять, почему они связаны в адронах, но и приводить к постоянному удержанию кварков внутри них.

Физиков не устраивал также тот факт, что в некоторых адронах составляющие их кварки, казалось бы, нарушали фундаментальный принцип квантовой механики — принцип запрета Паули, который применим ко многим частицам, включая кварки. Он гласит,

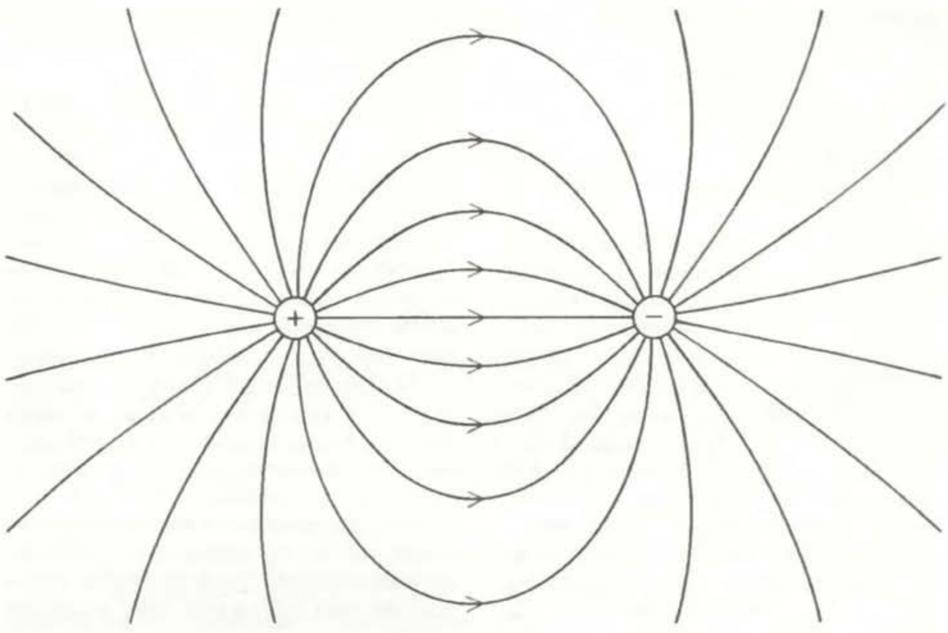
* Иногда *b*-кварк называют «прелестным» (от англ. beauty — красота, прелесть). — *Прим. перев.*

что две такие частицы, находящиеся в малой области пространства, не могут быть одновременно в одном и том же квантовомеханическом состоянии *. Это означает, что адрон не может состоять из двух одинаковых кварков, например из двух *u*-кварков, если направления их спинов не противоположны. Спин кварка подобен угловому моменту Земли, связанному с ее вращением. Однако в отличие от него спин кварка квантуется: он может принимать только одно из двух допустимых значений. По этой причине в любой группе из трех кварков по крайней мере два имеют одинаково направленные спины.

Существуют адроны, в которых на достаточно близком расстоянии должны находиться три одинаковых кварка. Частица Ω^- — один из таких адронов — была предсказана на основе кварковой модели, и ее последующее открытие в 1964 г. в Брукхейвене Н. Самиосом и Р. Шаттом с сотрудниками послужило убедительным подтверждением кварковой модели. Однако Ω^- — совершенно загадочная частица, поскольку, согласно кварковой модели, она должна состоять из трех *s*-кварков. Близкое расположение кварков в этом адроне, казалось, нарушает принцип Паули. По этой и другим причинам физики сначала предпочитали считать кварки просто удобной математической абстракцией. Вопрос об их физическом существовании был временно отложен.

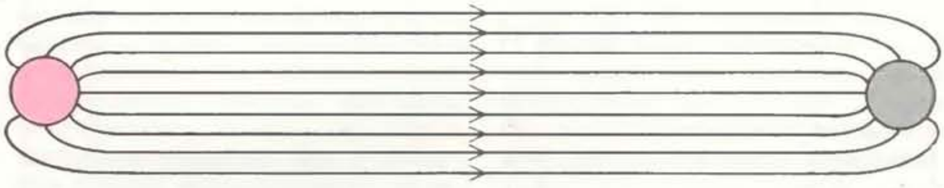
Примирение между кварковой моделью и принципом Паули произошло на основе идей, выдвинутых О. Гринбергом из Мэрилендского университета в Колледж-Парк и независимо Му-Юнг Ханом из Университета Дюка и Й. Намбу из Чикагского университета. Было сделано предположение, что кварк любого типа может находиться в трех разных состояниях. Например, если составить Ω^--частицу из *s*-кварка в состоянии *A*, *s*-кварка в состоянии *B* и *s*-кварка в состоянии *C*, то принцип Паули будет спасен. Чтобы как-то обозначить разные состояния кварков, физики стали приписывать им разные цвета. Кварк может быть красным, фиолетовым и зеленым, а антикварки окрашены в дополнительные цвета — голубой («антикрасный»), желтый («антифиолетовый») и малиновый («антизеленый»). Приставка «хромо» в названии «квантовая хромодинамика» как раз имеет отношение к цветовой терминологии.

Для правильной классификации адронов введение цвета должно быть дополнено еще одной гипотезой. Ведь



СИЛОВЫЕ ЛИНИИ электромагнитного поля расходятся в пространстве. Изображено поле двух частиц с противоположными электрическими зарядами. Гуще всего силовые линии расположены в пространстве между частицами, хотя они распространяются и в других направлениях. Напряженность поля в любой точке — сила, которая действует на частицу с единичным зарядом в этой точке. Если через данную точку провести поверхность единичной площади, перпендикулярную направлению силовых линий, то число силовых линий, пересекающих эту поверхность, будет пропорционально напряженности поля в этой точке. Напряженность электромагнитного поля точечного электрического заряда уменьшается как квадрат расстояния от него.

новая «цветовая» степень свободы наряду с возможностью построения кварковой модели таких частиц, как Ω^-, ведет к увеличению числа адронов. Например, Λ-частица состоит из *u*-, *d*- и *s*-кварков, и если каждый кварк может находиться в одном из трех состояний, то, казалось бы, должна быть не одна Λ-частица, как следует из наблюдений, а девять — по числу возможных комбинаций цветов. Чтобы избежать такого изобилия, вводится гипотеза, что кварки в адронах могут иметь только такую комбинацию цветов, которая оставляет сам адрон бесцветным, или белым, если предположить обычное, такое же, как для света, правило смешения цветов. Среди трех кварков в протоне или Λ-частице должны присутствовать один красный, один фиолетовый и один зеленый. Кварк и антикварк в π^+-мезоне могут быть красным и голубым, фиолетовым и желтым или зеленым и малиновым. По-



СЖАТИЕ силовых линий для системы из двух частиц в тонкую трубку постоянного сечения приводит к тому, что сила, действующая между ними, будет постоянна, независимо от расстояния между частицами. Тогда поверхность единичной площади, перпендикулярная продольной оси трубки, будет пронизываться одинаковым числом силовых линий, независимо от того, в каком месте вдоль трубки она расположена. Поскольку сила, удерживающая частицы вместе, постоянна, энергия для дополнительного разнесения частиц на фиксированное расстояние тоже не меняется, как бы далеко они ни находятся друг от друга. Необходима бесконечно большая энергия, чтобы «освободить» частицы. Таким образом, сжатие силовых линий приведет к постоянному удержанию частиц вместе. Если радиус трубки бесконечно мал, то пучок силовых линий подобен струне. Такая модель удержания кварков была математически сформулирована в 1968 г. Г. Венециано из Европейской организации ядерных исследований (ЦЕРН) и интерпретирована как струна Й. Намбу из Чикагского университета. Идея о физической реализации струны как пучка силовых линий была предложена в 1973 г. Х. Нильсеном и П. Олесеном из Института Нильса Бора в Копенгагене.

* Здесь имеются в виду частицы с полуцелым спином, подчиняющиеся статистике Ферми — Дирака. Кварк имеет спин 1/2 и, следовательно, относится к фермионам. — *Прим. перев.*

скольку «результирующий» цвет всегда тот же и в квантовомеханическом смысле каждое бесцветное состояние может осуществляться с одинаковой вероятностью, на самом деле существуют только одна Λ-частица и только один положительный π^+-мезон.

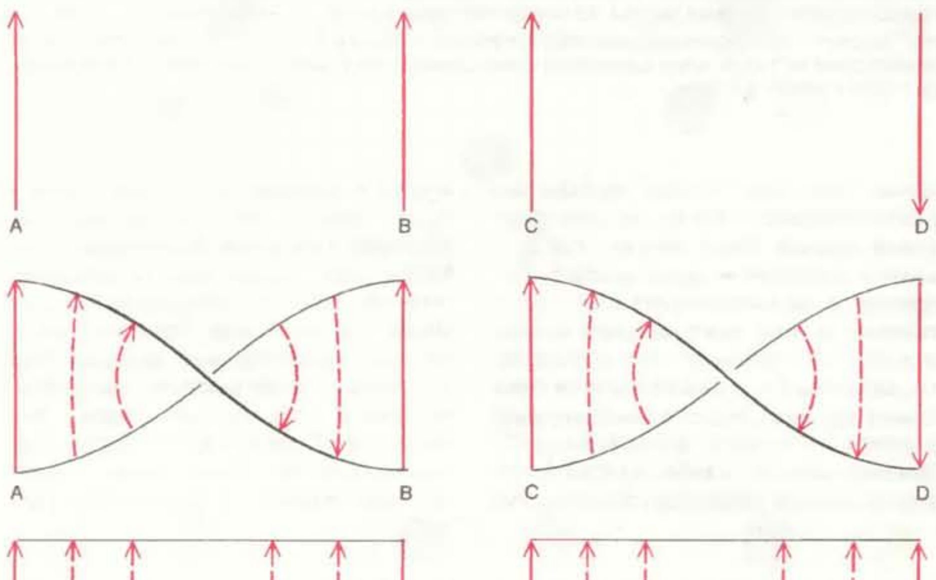
В КОНЦЕ 60-х годов на основе многочисленных экспериментов было получено убедительное доказательство того, что кварки в адронах — реальные частицы, а не чисто математические объекты. Исключительное значение имели серии экспериментов, выполненных на линейном ускорителе в Станфорде Дж. Фридманом и Г. Кендаллом из Массачусетского технологического института и Р. Тейлором из Ускорительного центра Станфордского университета (SLAC). Электроны высоких энергий направляют на протонную мишень, для того чтобы «прощупать» внутреннюю структуру протонов. В результате изучения частиц, рождавшихся в этих столкновениях, удалось показать, что в протоне имеются составляющие со всеми свойствами, приписываемыми кваркам.

Более того, хотя свободные кварки не были зарегистрированы, эксперименты показали, что в протоне они находятся в состоянии почти свободного движения. Это был совершенно удивительный результат. Как могут силы, способные постоянно связывать кварки между собой, позволить им в то же время двигаться почти свободно, когда они находятся столь близко друг к другу внутри протона? Можно представить, что три кваркоподобных объекта связаны между собой вроде трех камней, соединенных вместе в бóле — охотничьем оружии южноамериканских индейцев. Камни в боле движутся свободно, но это движение ограничено длиной соединяющего их ремня, который не дает им разлететься в разные стороны.

Новое экспериментальное свидетельство в пользу существования кварков вместе с идеей цвета дало сильный толчок к формулировке теории, описывающей динамику кварков. Цвет может служить источником поля, называемого хромоэлектрическим, которое приводит к взаимодействию нового типа между цветными частицами. В 1973 г. Д. Политцер из Калифорнийского технологического института и независимо Д. Гросс из Принстонского университета и Ф. Вильчек из Калифорнийского университета в Санта-Барбаре доказали, что динамическое взаимодействие, обусловленное хромоэлектрическим полем, должно привести к ослаблению сил между кварками при их сближении. Это объяснило почти свободное движение кварков внутри протонов, наблюдавшееся в станфордских экспериментах. Затем было сделано предположение, что то же взаимодействие может быть ответственно за удержание кварков, хотя тогда не было получено теоретических результатов в пользу столь привлекательной идеи. Тем не менее, коль скоро введено хромоэлектрическое поле, довольно произвольное предположение о бесцветности адронов и наблюдаемое удержание кварков могут рассматриваться как два аспекта одного и того же явления. Если какой-то из кварков, скажем красный, «вытащить» из адрона, то и сам кварк, и оставшаяся часть адрона будут цветными. Если цвет, подобно электрическому заряду, является источником поля, то между двумя цветными осколками могут появиться силы притяжения. Удержание кварков возникает, когда притяжение между двумя осколками настолько сильно, что их невозможно раздвинуть дальше определенного предела.

Чтобы объяснить некоторые свойства адронов, которые не могли быть объяснены существованием кварков, в начале 70-х годов была предложена динамическая модель, согласно которой адрон не является точечной или сферической частицей, а скорее может быть представлен в виде струны. Струна может вращаться или колебаться в соответствии с законами релятивистской динамики, а ее концы должны двигаться со скоростью света. Вычисления показали, что сила, действующая вдоль струны, должна быть огромна, приблизительно 14 т. Квантование колебаний натянутой струны приводит к различным состояниям, которым можно сопоставлять определенные адроны.

Очевидно, такую модель адронов можно объединить с их представлением в виде болы. Если три кварка или кварк и антикварк поместить на концы релятивистской струны, то тот факт, что они всегда связаны, можно объяснить ее натяжением. Однако модель струны, как и первоначальные представления о кварках, является по существу чисто математической абстракцией: струна — это одномерный объект.

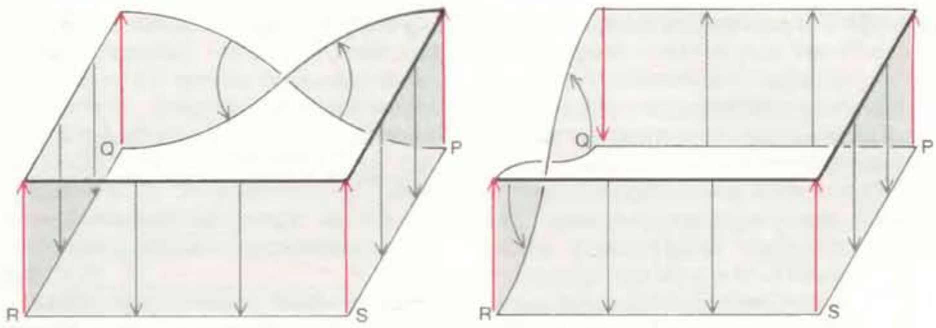


НАПРАВЛЕНИЯ ДВУХ СТРЕЛОК, помещенных в узлах решетки, нельзя сравнивать, пока не указан способ перемещения стрелок из одного узла в другой. На схеме вверху стрелки в узлах A и B направлены в одну сторону, в то время как в узлах C и D — в противоположные стороны. Такое сравнение, однако, заранее предполагает наличие листа бумаги, на которой нарисована схема. Если предположить, что каждая пара узлов соединена перекрученной бумажной лентой, как показано на схеме в середине, то сравнение направлений стрелок даст обратный результат. Без калибровочного поля, определяющего правила переноса стрелки из точки в точку, направления стрелок в соседних точках невозможно сравнивать. Схема внизу показывает, что если направление стрелки в любой точке изменить на противоположное, то результат сравнения не обязательно изменится. Он останется прежним, если при этом ленты соответственно перекрутить или раскрутить. Свойство калибровочного поля, дающее возможность скомпенсировать изменение направления каждой стрелки, называют локальной калибровочной инвариантностью.

Может ли она, несмотря на это, служить неким приближением к какой-то другой структуре, более приемлемой с физической точки зрения?

В 1973 г. Х. Нильсен и П. Олесен из Института Нильса Бора в Копенгагене указали, что струну можно интерпретировать как пучок силовых линий соответствующего поля. Силовые линии электромагнитного поля имеют хорошо известную всем конфигурацию. Напряженность поля пропорциональна плотности силовых линий. Таким образом, когда силовые линии расходятся, как это происходит при удалении от полюсов обычного магнита, напряженность поля убывает. Если, однако, сжать силовые линии в «трубку» постоянного сечения, то напряженность поля будет постоянной вдоль всей трубки. Сила, необходимая для «растаскивания» кварка и антикварка на противоположные концы такой трубки, также должна быть постоянной независимо от того, как далеко находятся друг от друга частицы. Тогда, чтобы высвободить один из кварков, потребуется бесконечно большая энергия.

Квантовомеханическая природа кварков и струн требует, чтобы силовые линии цветового взаимодействия кварков вели себя совершенно иначе, чем силовые линии взаимодействия электрически заряженных частиц. Поскольку оба взаимодействия распространяются в вакууме, можно предположить, что любое различие между ними возникает вследствие их внутренней природы, а не из-за влияния вакуума. В классической, ньютоновской, механике такое предположение было бы логичным. Действительно, поскольку классический вакуум, по определению, есть состояние без вещества и энергии, между ним и полем не может быть никакого взаимодействия. Но в рамках квантовой механики даже вакуум имеет свою структуру, которая может изменить распространение полей и сил.



В РАЗУПОРЯДОЧЕННОЙ ГРАНИ одна из лент, образующих ее стороны, перекручена. Ленты изображают калибровочные поля. Чтобы компенсировать изменение направления стрелки в любой точке, калибровочное поле может быть локально перекручено или раскручено, однако полностью снять перекручивание во всех четырех лентах невозможно. Например, ленту, перекрученную между точками P и Q (*слева*), можно раскрутить, поменяв направление стрелки в Q на противоположное, тогда перекручивание перейдет на ленту между Q и R (*справа*). Независимо от того, какие стрелки изменили направление, т.е. сколько локальных калибровочных преобразований было совершено, нечетное число лент из четырех должно быть перекручено. В результате грань останется разупорядоченной, несмотря на локальную калибровочную инвариантность.

Структура вакуума является следствием принципа неопределенности Гейзенберга. В одной из формулировок этот принцип гласит, что в любом физическом событии наблюдаемая энергия связана с неопределенностью во времени, в течение которого происходит это событие. Более точно, произведение неопределенности в энергии на неопределенность во времени не может быть меньше некоторой числовой постоянной. Для события, ограниченного рамками очень короткого интервала времени, имеется соответственно бо́льшая неопределенность в энергии. Поэтому в квантовом вакууме для любого короткого интервала времени существует вероятность обнаружить отличную от нуля энергию.

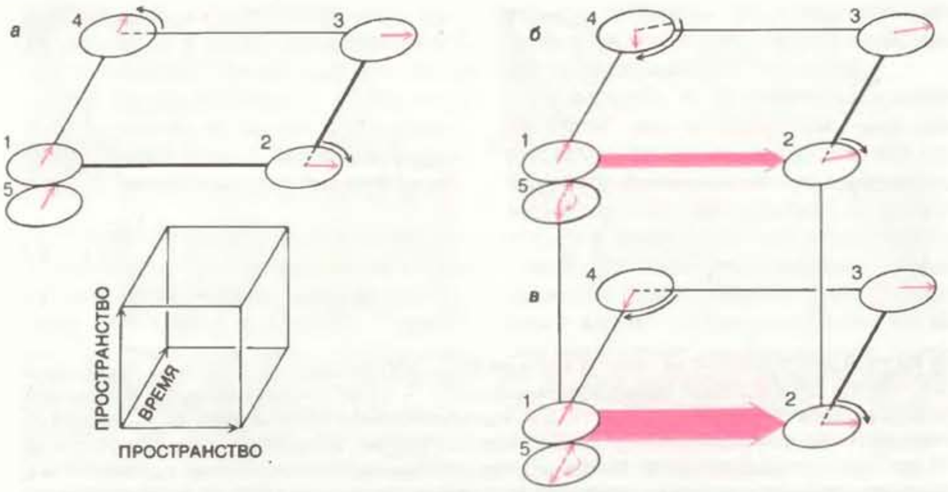
Энергия вакуума может проявить себя в спонтанном рождении или уничтожении частиц и их античастиц либо в появлении и исчезновении электрического или хромоэлектрического поля в различных областях пространства. Такие изменения квантового поля называются флуктуациями. Например, квантовые флуктуации электромагнитного поля между двумя электрически заряженными частицами приводят к тому, что взаимодействие зарядов не определяется полностью классическим полем, которое описывается уравнениями Максвелла. Измеряемое электромагнитное поле является средним всех полей, возбужденных квантовыми флуктуациями. Усреднение производится с весом, определяемым вероятностью появления данной флуктуации.

В большинстве практических приложений электродинамики эффекты квантовых флуктуаций весьма малы. Измерение поля между двумя макроскопическими электрически заряженными телами согласуется с значением, предсказываемым классической теорией. Однако при столкновении заряженных частиц высоких энергий квантовые флуктуации становятся гораздо более существенными и должны приниматься во внимание при расчетах электромагнитных эффектов. При стандартном подходе, очень успешном для решения многих задач, сначала вычисляют свойства поля в классическом вакууме. Затем результат классических вычислений последовательно уточняется с учетом квантовых флуктуаций. Такую процедуру называют разложением в ряд теории возмущений. В квантовой теории электромагнетизма чем больше или сложнее флуктуация, тем меньше вероятность ее появления. Поэтому почти все поправки к классическому электромагнитному полю, которые должны быть учтены при квантовомеханических расчетах, обусловлены малыми флуктуациями.

Можно предположить, что свойства хромоэлектрического поля между двумя кварками выводятся аналогичным образом и что в принципе разложение в ряд теории возмущений позволяет найти напряженность поля в любой точке с любой желаемой степенью точности. Однако методы теории возмущений применимы только в том случае, когда поле, рассчитанное для классического вакуума, определяет основной эффект, т.е. когда поправки на флуктуации малы и уменьшаются при рассмотрении более сильных флуктуаций. Для квантовомеханических явлений, которые главным образом зависят от вклада больших флуктуаций, ряд теории возмущений не сходится. Говорят, что такие явления выходят за рамки теории возмущений. Среди них — «запирание» в трубку силовых линий хромоэлектрического поля и, следовательно, постоянная связь двух кварков между собой.

Как можно продемонстрировать явление удержания кварков? Теоретики поняли, что необходим новый подход, позволяющий включать в рассмотрение большие квантовые флуктуации уже в самом начале вычислений. Таким подходом и стал метод решетки, предложенный Вильсоном.

Решетка обычно является кубической, ее можно представить как совокупность вершин и ребер множества плотно сложенных кубиков. Она простирается как в пространстве, так и во времени, так что каждая ее точка указывает и положение в пространстве, и момент времени. Для наглядности решетку можно представить как упорядоченное множество кубиков, в которых две оси координат относятся к пространственным переменным, а третья — к временно́й переменной. Настоящая решетка, разумеется, трех-



ИДЕЮ КАЛИБРОВОЧНОГО ПОЛЯ можно обобщить на случай, когда величины, определенные в каждой точке решетки, изменяются непрерывно. Например, можно направить стрелку в каждой точке под произвольным углом к вертикали. Калибровочное поле позволяет сравнивать углы в разных точках решетки. Если после полного обхода грани стрелка не возвращается к первоначальной ориентации, угол между ее начальным и конечным положениями (который называют фазовым углом) есть мера разупорядоченности грани. Таким калибровочным полем является электрическое поле. Проекция напряженности поля в данной точке на определенное направление определяется степенью разупорядоченности соответствующей грани. Рисунок иллюстрирует, как разупорядоченность грани связана с отдельной компонентой электрического поля, а именно с пространственной компонентой, направленной слева направо. Строго говоря, решетка имеет четыре измерения — три пространственных и одно временнóе. Но ее можно представить с помощью двух пространственных и одной временнóй координат. Когда лимб со стрелкой обносят вокруг грани, имеющей направленную слева направо пространственную компоненту, и временнýю компоненту, калибровочное поле заставляет стрелку на лимбе вращаться. Если при полном обороте лимба вокруг грани стрелка возвращается в первоначальное положение, то напряженность поля в этом направлении равна нулю (*а*). С увеличением фазового угла растет и пространственная компонента напряженности поля (*б* — *в*); это показано увеличением толщины цветной стрелки, указывающей направление поля. С математической точки зрения вращение стрелки по лимбу можно отождествить с ее вращением в плоскости комплексных чисел.

мерна в пространстве и также имеет временнýю ось, являясь, таким образом, четырехмерной структурой. Между любыми двумя узлами решетки имеется ребро, которое можно изобразить линией, соединяющей их. Квадрат, ограниченный четырьмя ребрами, называют гранью. В формулировке Вильсона узлы, ребра и грани — это все, что остается от обычного пространства и времени. Ребра и грани должны рассматриваться как еще более абстрактные объекты, чем узлы решетки, которые и являются единственными ее точками. Ребра и грани полностью определяются узлами решетки и других точек помимо узлов не содержат. Другими словами, пространство и время на решетке совсем не похожи на обычные пространство и время, в которых между любыми двумя заданными точками имеется бесконечное множество других.

Введение пространственно-временной решетки (по Вильсону) не означает, разумеется, что физические процессы действительно происходят в дискретных точках на решетке. Сейчас общепринято, что пространство-время непрерывно. Решетка представляет собой всего лишь то, что в теоретической физике называют регуляризацией — промежуточной процедурой, введенной для проведения вычислений, которые иначе были бы невыполнимы. В применении к проблеме удержания кварков такой подход состоит в следующем. Все частицы определены на узлах решетки, а напряженность поля — вдоль ее ребер. (В каждом узле задается лишь вероятность обнаружения там частицы. Из самой идеи решетки ясно, что вероятность нахождения частицы между двумя соседними узлами не определена, т.е. не имеет смысла.) В отсутствие частиц симметрия флуктуаций на решетке в сильной степени упрощает вычисления среднего электрического или хромоэлектрического поля, возбуждаемого большими флуктуациями. Поля характеризуются не только величиной, но и направлением: с одинаковой вероятностью они могут быть направлены как в одну сторону вдоль ребра решетки, так и в противоположную. Таким образом, в вакуумном состоянии в отсутствие частиц среднее значение электрического или хромоэлектрического поля по всей решетке равно нулю.

Аналогичные, но более сложные вычисления можно провести для больших флуктуаций поля в том случае, когда на решетке задана одна частица и соответствующая ей античастица. В среднем флуктуации поля по-прежнему сокращаются, за исключением тех, которые заданы вдоль ребер решетки, образующих кратчайший путь между частицей и античастицей. Результаты не зависят от типа поля, заданного на решетке: частицы могут быть кварком и его антикварком или электроном и позитроном. Таким образом, свойство удержания — естественное следствие определения полей на решетке.

Следующий шаг — снять решетку и вернуться в обычное непрерывное пространство-время. Для этого нужно уменьшать постоянную решетки (расстояние между соседними узлами) таким образом, чтобы ее узлы располагались все плотнее в пространстве-времени. Если уменьшение постоянной решетки имеет математический предел, то восстановится непрерывное пространство-время. В пределе эта процедура позволяет вычислить среднее поле с учетом уже всех квантовых флуктуаций.

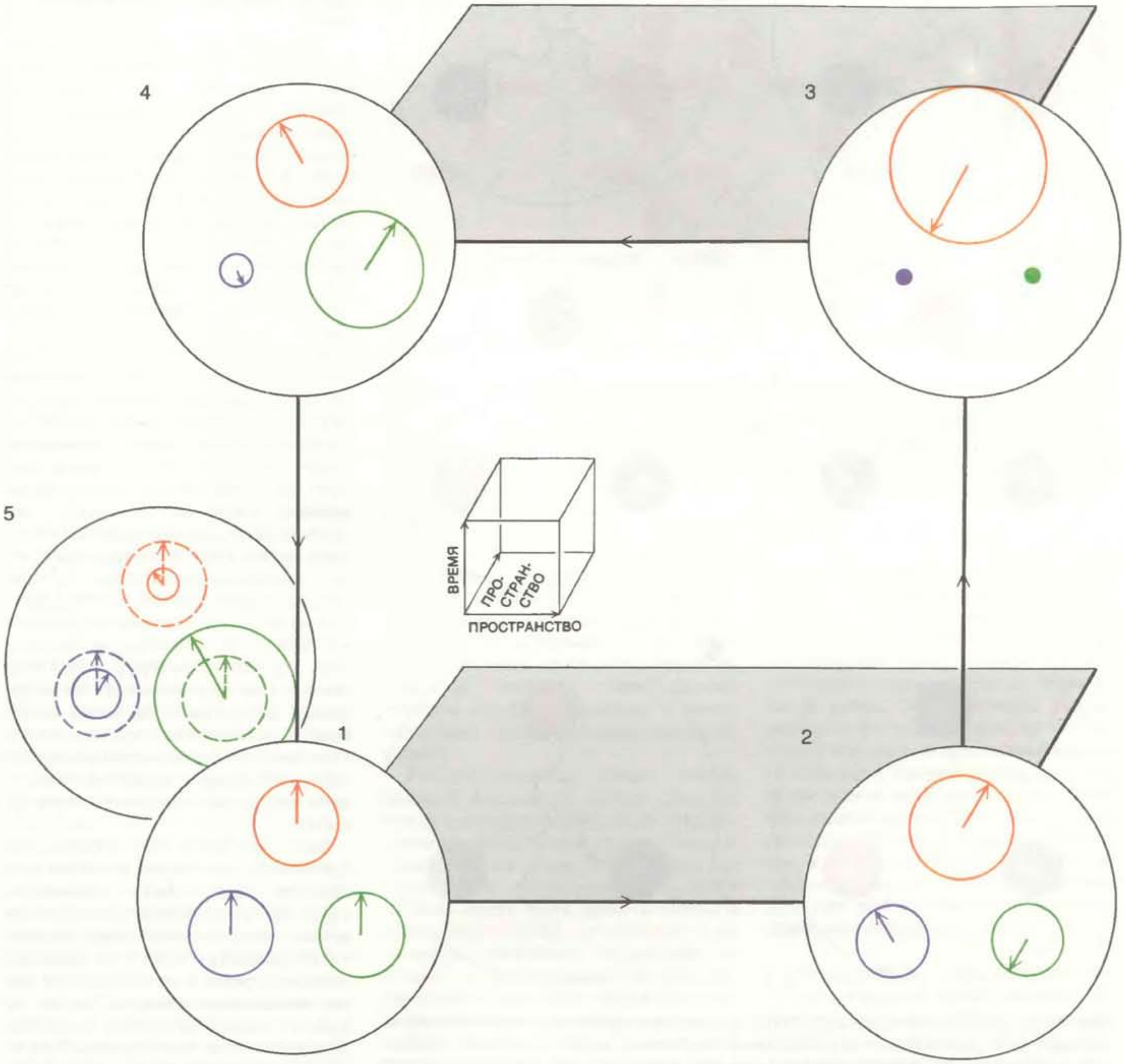
ЧТО ЖЕ достигается введением решетки? Большие флуктуации, ответственные за сжатие цветовых силовых линий, сначала направлены вдоль каждого из ребер решетки. Однако пока ячейки решетки остаются относительно крупными, квантовые флуктуации ведут к запиранию как хромоэлектрического, так и обычного электромагнитного поля.

Хорошо известно, что в отличие от цветовых зарядов электрические заряды могут существовать изолированно. Поскольку метод решетки предсказывает удержание и электрического заряда, такой подход должен вызывать сомнения, пока не будет показано, что при уменьшении ячеек решетки электромагнитное поле подчиняется хорошо установленным законам. Другими словами, хотелось бы показать, что на некоторой стадии сжатия решетки электромагнитные силовые линии должны «освобождаться из заключения» на ребрах решетки, в то время как при переходе к непрерывному пространству-времени хромоэлектрические силовые линии все время остаются сжатыми в пучок. Доказать, что все происходит именно так, довольно сложно, но за последние несколько лет это было сделано с помощью высокоскоростных компьютеров.

Фундаментальные различия между электрическим и хромоэлектрическим полями физики-теоретики связывают с тем, что КЭД является абелевой калибровочной теорией, в то время как КХД — неабелева калибровочная теория поля. Эта терминология связана с

именем норвежского математика Нильса Абеля. Различие абелевых и неабелевых теорий выводится из математической теории групп, описывающей симметрию, присущую последовательности таких операций, как вращения. Если операции, принадлежащие какой-то группе, могут быть выполнены в любой последовательности с одинаковым конечным результатом, то группа является абелевой. Например, группа вращений относительно какой-нибудь оси абелева, поскольку такие вращения дают один и тот же эффект независимо от их последовательности. Если же порядок, в котором выполняются две или более операций, влияет на конечный результат, то группа операций будет неабелевой. Вращения кубика вокруг трех его осей образуют неабелеву группу: когда кубик вращается вокруг вертикальной и горизонтальной осей, окончательный результат зависит от того, какая операция выполнялась первой.

Чтобы представить, каким образом теория групп применяется к КХД и КЭД, необходимо понять основную идею калибровочного поля. Ее можно проиллюстрировать на примере изолированных точек пространства и времени, расположенных на решетке. Частицы могут находиться в одном из узлов решетки или «перепрыгивать» из одного узла в другой; при движении их состояние может меняться. Состояние частицы определяется набором величин, значения которых могут изменяться в некотором интервале. Для упрощения примем, что состояние частицы описывается одной переменной, которая принимает только два значения. Например, в каждом узле решетки может быть задана переменная со значением либо +1, либо −1, указыва-



ХРОМОЭЛЕКТРИЧЕСКОЕ ПОЛЕ является калибровочным полем, в принципе похожим на электромагнитное, но его математическое описание более сложно. В каждой точке решетки вместо одной стрелки имеются три: они соответствуют цветовым зарядам кварков. Более того, цветовое калибровочное поле влияет не только на направление каждой стрелки, но и на ее длину. Длины стрелок не являются независимыми: корень квадратный из суммы квадратов их длин должен быть равен 1. Напряженность хромоэлектрического поля вдоль каждого ребра решетки зависит от фазовых углов и изменения конфигурации трех стрелок после полного оборота вокруг грани.

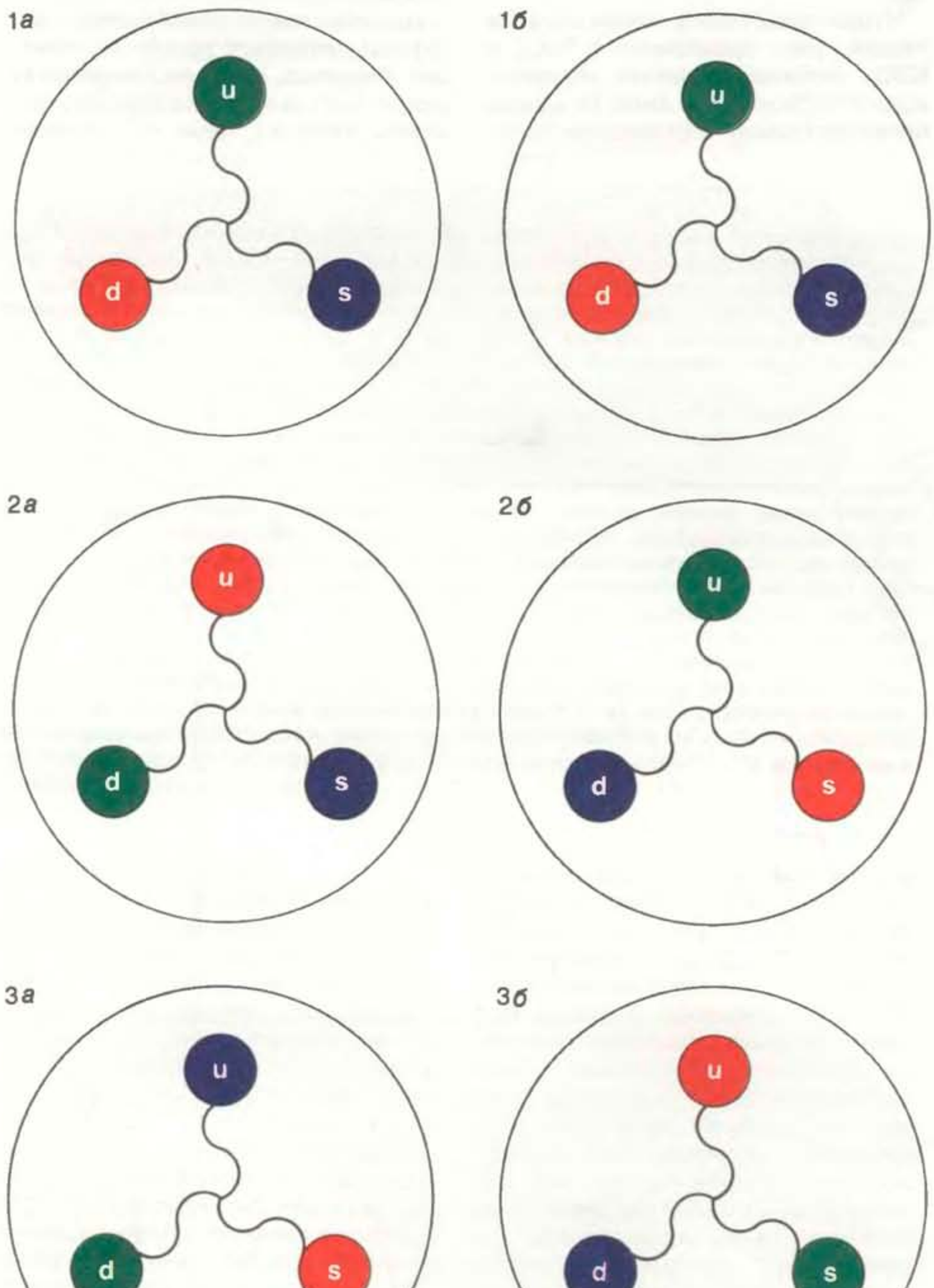
ющим знак электрического заряда; изобразим их стрелкой, направленной соответственно либо вверх, либо вниз.

При описании процесса распространения взаимодействия в пространстве-времени важно иметь возможность сравнивать значения переменных в соседних точках. Для сравнения длины двух предметов, находящихся в удаленных областях пространства, необходима измерительная линейка или калибр, которые можно поднести к одному из предметов, сделать отметки и перенести к другому предмету, а затем сравнить их длины. Аналогично на решетке надо иметь возможность сравнивать ориентацию стрелок в соседних узлах. На первый взгляд это кажется тривиальной задачей. Допустим, что на листе бумаги расположены два узла и в каждом из узлов имеется направленная вверх стрелка. Разве не очевидно, что обе стрелки указывают одно направление? Однако мы заранее предположили, что существует лист бумаги, на котором нарисована решетка. Бумага здесь выступает в роли посредника, позволяющего сравнить обе ориентации. Мы как бы мысленно переносим стрелку из одного узла в другой и приходим к выводу, что срелки имели одинаковую ориентацию до переноса, поскольку после переноса их направление совпало.

Предположим теперь, что лист бумаги между двумя стрелками перекручен (см. рисунок на с. 40). Стрелка, указывающая вверх, после переноса по перекрученному листу во второй узел станет указывать вниз. Так как направление стрелки определено лишь в изолированных точках решетки, невозможно установить, какой из двух методов переноса правилен. Действительно, без листа бумаги, задающего правила переноса стрелки, невозможно сравнить ее направления в различных узлах.

Калибровка в теории поля есть некий эталон, аналогичный расстоянию между двумя отметками на металлическом бруске или направлению стрелки на лимбе (круговой шкале). Калибровка может меняться под влиянием поля при перемещениях в пространстве-времени. Поле, которое может вызвать такие изменения, называют калибровочным; оно точно указывает, какие предположения сделаны относительно правил переноса калибровки. В нашем простом примере калибровочное поле — это набор правил для переноса стрелок вдоль ребер решетки из одного узла в следующий. Калибровочное поле можно представить как бумажную ленту, соединяющую соседние узлы решетки и позволяющую переносить и сравнивать взаимную ориентацию стрелок из различных точек решетки.

Важно отметить, что стрелка, изображающая состояние в любом узле решетки, может быть перевернута, тогда лента, изображающая калибровочное поле, соответственно перекрутится или раскрутится. Если переворачивание стрелки и перекручивание ленты согласованы между собой, то результат сравнения любых состояний до и после таких преобразований не изменится. Стало быть, не изменится физическая информация о рассматриваемой системе. Возможность переворачивания стрелок без изменения физической информации называют локальной калибровочной инвариантностью.



ГРУППА ЦВЕТОВЫХ ПРЕОБРАЗОВАНИЙ в хромоэлектрическом поле с математической точки зрения является неабелевой. Неабелева группа — абстрактный набор таких операций (вместе с объектами, над которыми они совершаются), что конечный результат нескольких операций зависит от последовательности их выполнения. В хромоэлектрическом поле к таким операциям относится обмен цветом между кварками. На рисунке изображены два разных способа обмена цветом для данной кварковой конфигурации (1*а*, 1*б*). На рисунке слева кварки сначала обмениваются красным и зеленым цветами (2*а*), а затем фиолетовым и красным (3*а*). Порядок обмена на рисунке справа обратный (2*б*, 3*б*). Окончательные кварковые конфигурации различны.

Из-за локальной калибровочной инвариантности можно прийти к выводу, что калибровочное поле представляет собой излишнее усложнение. Каков смысл введения перекрученной ленты, если она потом может быть раскручена обратно с помощью локального калибровочного преобразования? Возражение справедливо до тех пор, пока рассматриваются только пары точек на решетке. Однако значение локальной калибровочной инвариантности становится очевидным, когда переходят к рассмотрению грани на решетке — совокупности четырех точек, расположенных в углах квадрата и соединенных вдоль ребер решетки четырьмя лентами, изображающими калибровочное поле. Представим такую грань, у которой одна из лент перекручена, а остальные три нет (см. рис. на с. 41). Единственную перекрученную ленту можно раскрутить обратно; при этом одна из стрелок перевернется. Но это приведет к перекручиванию одной из трех других лент. Независимо от того, какие стрелки переворачиваются и сколько локальных калибровочных преобразований выполняется, на такой грани всегда должно быть перекручено нечетное число лент. Грань, для которой невозможно раскрутить все четыре ленты, называют разупорядоченной гранью.

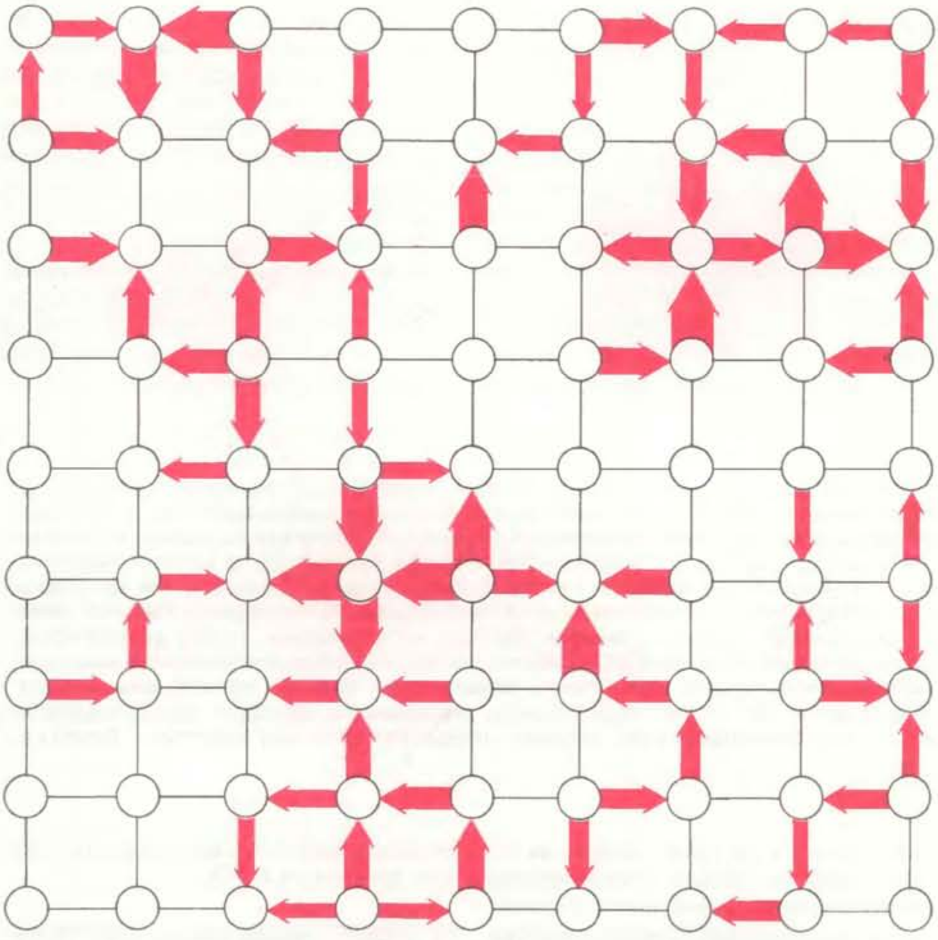
Разупорядоченность* грани проявляется и другим образом. Стрелка, направленная вверх, при переносе вдоль лент вокруг разупорядоченной грани вернется в исходный узел, указывая вниз. Таким образом, разупорядоченные грани на решетке — это такие грани, при переносе стрелки вокруг которых изменяется ее ориентация.

Понятие разупорядоченной грани нетрудно применить к обычному пространству-времени. В физическом пространстве-времени нельзя непосредственно увидеть ленты, вдоль которых перемещаются стрелки. Как бумажные ленты, изображающие калибровочное поле на решетке, не являются частью решетки, так и сами калибровочные поля в обычном пространстве-времени не являются его частью. С математической точки зрения абстрактное пространство высшей размерности, определяющее вращения стрелки, называют связностью в расслоенном пространстве (см. статью Г. Бернстейна и Э. Филлипса «Расслоения и квантовая теория», "Scientific American", июль, 1981)*. Даже без полного понимания математического описания можно представить, что стрелка, движущаяся вдоль некоторого замкнутого пути в пространстве-времени, способна вернуться в исходную точку с изменившимся направлением.

Разупорядоченные грани имеют важный физический смысл. Энергия любого калибровочного поля сосредоточена именно на этих гранях. Неразупорядоченные грани, для которых все ленты могут быть раскручены и все стрелки могут быть ориентированы в одном направлении, соответствуют вакуумному состоянию физической системы, т.е. конфигурации без энергии. На разупорядоченных гранях осуществляются квантовые флуктуации.

Идея калибровочного поля может быть легко обобщена на более сложные случаи. Например, стрелкам можно позволить образовывать любой угол по отношению к некоторому фиксированному направлению, а не ограничиваться направлениями «вверх» и «вниз». Тогда калибровочное поле будет определять угол, на который поворачивается стрелка при ее переносе вдоль ребра. Соответственно разупорядоченность грани может проявиться в том, что угол между направлениями стрелки до и после ее переноса вокруг грани может принимать любое значение в непрерывном интервале. Степень разупорядоченности грани, выраженная в соответствующих математических единицах, называют ее действием. Действие всей системы есть сумма действий всех отдельных граней.

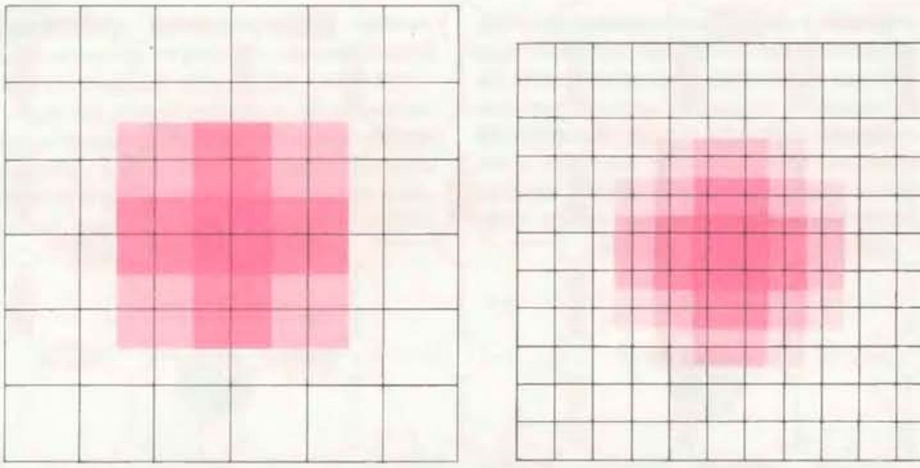
Оказывается, что электрическое поле является калибровочным полем, определяющим непрерывные вращения стрелки вокруг одной оси, хотя это становится очевидным лишь в наиболее общей и усложненной формулировке электромагнетизма. Более распространены представления об электрическом поле как о наборе векторов, причем в каждой точке пространства вектор определяет величину и направ-



ФЛУКТУАЦИИ СИЛОВЫХ ЛИНИЙ НА РЕШЕТКЕ характеризуют квантовый вакуум. Флуктуации являются следствием принципа неопределенности Гейзенберга. В отсутствие заряженных частиц флуктуации компенсируют друг друга, поэтому измерение, скажем, электрического поля в вакууме дало бы нулевой результат, так как измерялось бы среднее значение всех флуктуаций. Тем не менее флуктуации силовых линий в вакууме должны учитываться при вычислении силового поля двух заряженных частиц.

* Этот термин (от англ. frustration) применяется в статистической теории разупорядоченных систем — так называемых спиновых стекол. В моделях таких систем рассматривается случайное распределение ферромагнитного и антиферромагнитного взаимодействия между спинами, находящимися в соседних узлах решетки. Графически характерная ситуация может быть изображена с помощью тех же лент и стрелок. — *Прим. перев.*

* Перевод на русский язык см.: журнал «Успехи физических наук», 1982, т. 136, с. 665—692. — *Прим. перев.*



СЖАТИЕ РЕШЕТКИ, на которой определены частицы и поля, необходимо для того, чтобы можно было воспроизвести эффекты, обусловленные полем в обычном, непрерывном, пространстве-времени. Однако в то время как постоянная решетки уменьшается, размер составной частицы, такой, как протон, не должен изменяться. Это означает, что вероятность обнаружения кварка в данном узле решетки в протоне должна быть отлична от нуля уже на большем числе узлов. Рисунок справа показывает результат сжатия решетки, изображенной слева, до размеров, вдвое меньших первоначальных. Интенсивность окраски показывает, насколько велика относительная вероятность обнаружения кварка в данном узле. Сжатие может быть обеспечено постепенным уменьшением значения так называемой константы взаимодействия, которая определяет величину квантовых флуктуаций.

ление поля в этой точке. Однако на основе наиболее общей формулировки калибровочной теории было установлено, что величина каждого вектора прямо пропорциональна разупорядоченности соответствующей грани.

Вращающейся стрелке электромагнитного калибровочного поля соответствует комплексное число, которое изменяет свое значение при переносе стрелки вокруг грани. (Стрелку и комплексное число не следует путать с вектором, представляющим само электрическое поле на каждом ребре решетки.) Грань — это петля в пространстве-времени, так что переносимую стрелку надо представлять движущейся вперед-назад как в пространстве, так и во времени. Стрелка перемещается, скажем, в положительном направлении вдоль какой-нибудь пространственной оси, например оси x, затем вперед во времени, потом назад по оси x и, наконец, назад во времени к исходной точке. Количественную характеристику изменения направления переносимой стрелки называют фазовым углом. Согласно калибровочной теории, величина x-компоненты вектора, обычно связываемого с напряженностью электрического поля в точке, является мерой фазового угла, образующегося при обходе вокруг пространственно-временной петли, начинающейся в направлении оси x. Поэтому квантовые флуктуации электрического поля в некоторой точке можно представлять как флуктуации величины фазового поворота стрелки, когда последняя переносится вокруг грани, простирающейся в одном из пространственных измерений и вдоль оси времени (см. рисунок на с. 42).

Трудно избавиться от некоторого ощущения нереальности, когда электрическое поле рассматривается как абстрактное пространство фазовых поворотов. Ведь электрическое поле вполне осязаемо, особенно если вас ударит током. Тем не менее по мере продвижения вглубь в изучении физического мира осязаемые факты и математические концепции переплетаются. Абстрактная идея фазового пространства в конечном итоге связана с различиями между квантовой электродинамикой и квантовой хромодинамикой, т.е. с различиями между абелевыми и неабелевыми калибровочными теориями. Основной групповой операцией в КЭД являются повороты фазовых углов. Результат двух последовательных фазовых поворотов, скажем на углы 30 и 50°, не зависит от порядка, в каком они выполняются. Окончательным результатом в любом случае будет поворот на угол 80°. Это характерно для абелевых групп, поэтому КЭД — абелева калибровочная теория.

Если переменными, заданными в узлах решетки, являются цвета кварков, переносимая калибровка состоит из трех стрелок вместо одной и стрелки могут изменять не только направление, но и свою длину (см. рисунок на с. 43). Каждая стрелка изображает один из трех цветовых зарядов; все вместе цветовые заряды определяют цвет кварка. При перемещении кварка вдоль ребра решетки он может изменить свой цвет с красного на зеленый; в результате красный кварк может поменяться цветом с зеленым кварком, а только что «позеленевший» кварк — с фиолетовым кварком. Однако окончательный результат таких обменов зависит от их последовательности. Поскольку это характерно для неабелевых групп, КХД — неабелева калибровочная теория.

Неабелева природа КХД вносит дополнительные степени свободы для флуктуаций хромоэлектрического поля. Более того, существование трех цветов приводит к тому, что грани для хромоэлектрического поля могут быть разупорядочены гораздо большим числом способов, чем грани для электромагнитного поля. По-видимому, дополнительные степени свободы и отсутствие упорядоченности в неабелевых теориях ответственны за удержание кварков в адронах.

Я уже говорил, что для грубых решеток удержание кварков можно доказать сравнительно простым расчетом больших квантовых флуктуаций. Для меньших флуктуаций на более мелких решетках вычисления становятся гораздо более сложными. Значение любой физической величины, измеряемой в эксперименте, является квантово-средним значением, т.е. взвешенным средним по всем возможным ее значениям. Измеряемое электрическое или хромоэлектрическое поле является средним по всем возможным конфигурациям поля, возбуждаемого флуктуациями, но не все конфигурации дают одинаковый вклад в среднее поле. Поэтому каждая из них должна быть взвешена, т.е. умножена на некоторый фактор, связанный с вероятностью возникновения данной конфигурации. В принципе, квантово-среднее физической величины, заданной на решетке, может быть вычислено в две стадии. Сначала определяют значение для каждой конфигурации поля и умножают его на соответствующий весовой множитель. Затем эти произведения складывают и полученный результат делят на сумму весовых множителей.

Однако даже для решеток, введенных в небольшом объеме пространства-времени, число возможных конфигураций настолько велико, что вопрос о точном суммировании даже не стоит. Для простейшей из возможных калибровочных теорий, когда калибровочное поле между двумя узлами решетки может быть либо перекручено, либо нет, число конфигураций для решетки из 10 узлов в каждом направлении составляет 2^{40000}, или больше чем 10^{12041}.

Поэтому для вычисления квантово-среднего значения поля на решетке необходимы методы статистической выборки. Аналогичные методы используют в США при выяснении обществен-

ного мнения. Чтобы определить, как спорный вопрос воспринимается различными группами населения, делается выборка, поскольку невозможно узнать мнение каждого человека. Результат выборочного опроса правильно отражает мнение всех групп в целом, если вероятность выбора числа отвечающих на вопрос из данной группы подобрана в соответствии с долей, или весом, этой группы во всем населении страны.

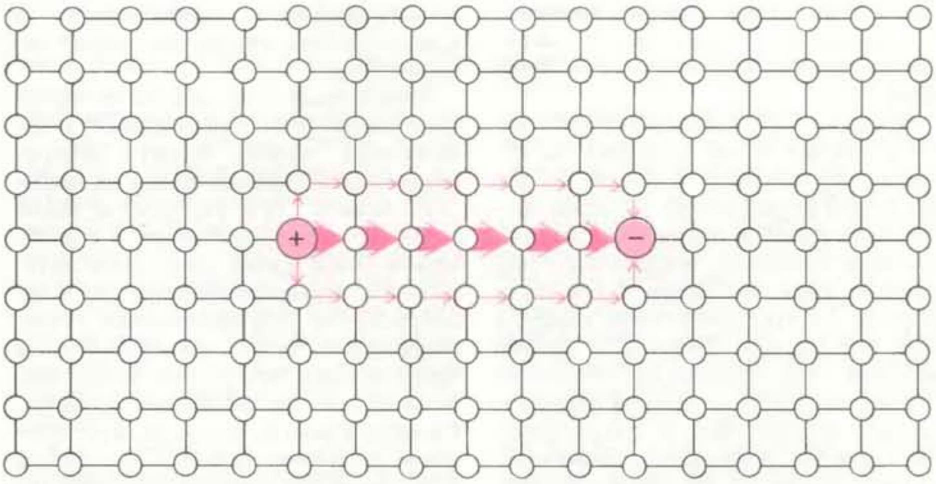
Аналогично программа компьютера выбирает определенные конфигурации квантовых флуктуаций. Компьютер воспроизводит огромное число конфигураций (но намного меньше, чем 10^{12041}); вероятность появления конкретной конфигурации полагается равной квантовому весовому множителю для этой конфигурации. Выборка конфигураций ведет к получению усредненного значения для квантового поля, которое получилось бы и при рассмотрении полного набора конфигураций.

Величина вклада данной конфигурации при вычислении квантово-среднего значения определяется ее действием, которое обычно обозначают S; поэтому весовой множитель дается некоторой математической функцией действия S. Изящная формула для весового множителя в квантовомеханических системах была выведена Р. Фейнманом из Калифорнийского технологического института. Сначала полное действие данной конфигурации делят на некоторую постоянную g^2, а затем вычисляют экспоненциальную функцию этой величины. Другими словами, число e (приблизительно равное 2,7) возводится в степень S/g^2. Весовой множитель обратно пропорционален полученному результату. Таким образом, весовой множитель пропорционален $1/\exp(S/g^2)$.

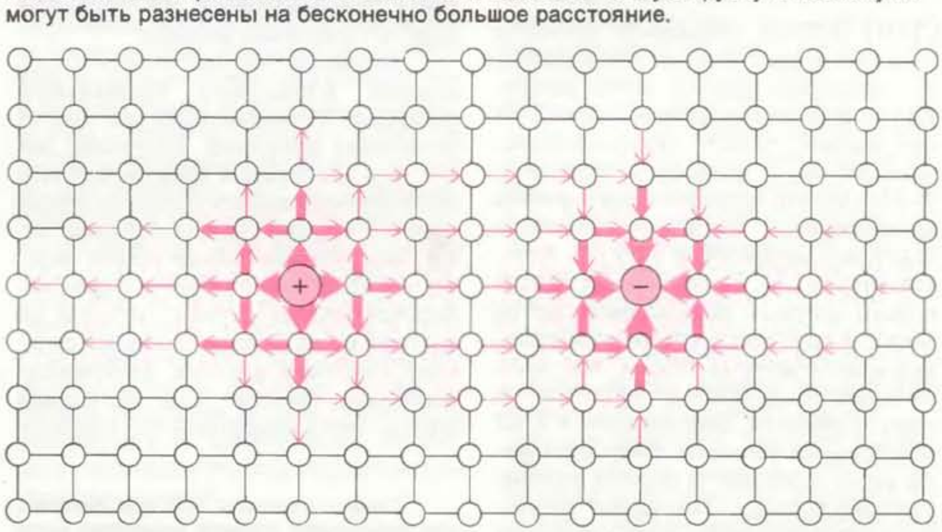
Из формулы для весового множителя следует, что чем больше действие конфигурации, тем меньше эта конфигурация весит при вычислении среднего квантового поля. Но в показатель экспоненты входит не само S, а S/g^2, поэтому изменение g может заметно влиять на весовой множитель данной конфигурации. Если g велико, то S/g^2 мало и величина $1/\exp(S/g^2)$ больше, чем она могла бы быть при малом g. Величину g называют константой связи. Таким образом, когда константа связи велика, весовой множитель квантовой флуктуации с большим действием выше, чем для малой константы связи.

Константу связи можно определить как внутреннюю меру силы физического поля. В теории электромагнетизма эта величина играет важную роль; ее значение примерно равно 1/137. Однако довольно абстрактным свойством калибровочных теорий на решетке является то, что константу связи g следует понимать как величину, принимающую фиксированное значение при заданной постоянной решетки, но которая может меняться вместе с изменением последней. Доказав удержание кварков на грубой решетке, мы должны уменьшать ячейку решетки, постепенно снижая значение константы связи. Процесс, при котором решетка делается все мельче и мельче, пока не будет достигнут предел непрерывного пространства-времени, называют перенормировкой.

Строго говоря, пока еще невозможно перенормировать решетку для флуктуаций хромоэлектрического поля, т.е. устремить постоянную решетки к нулю. Тем не менее можно уменьшать постоянную решетки, оставляя ее все же больше нуля, и искать указания на то, что силовые линии остаются (или не остаются) «запертыми» на более мелких решетках. В то время как постоянная решетки уменьшается, размеры всех физических объектов должны оставаться неизменными. На грубой решетке распределение вероятностей для протона может быть выбрано так,



«ЗАПИРАНИЕ» ПОЛЯ, будь то электрическое или хромоэлектрическое поле, можно продемонстрировать достаточно простыми расчетами в том случае, если оно определено на грубой решетке и флуктуации силовых линий велики. Значительные по величине флуктуации гасят друг друга везде, кроме области между заряженными частицами. Силовые линии сжимаются в тонкую трубку и поэтому не могут быть разнесены на бесконечно большое расстояние.



ВЫЧИСЛЕНИЯ ЭЛЕКТРИЧЕСКОГО ПОЛЯ двух противоположно заряженных частиц на решетке были выполнены автором с помощью компьютера CDC-7600 (Control Data Corporation) в Брукхейвенской национальной лаборатории. На рисунке продемонстрирован результат нескольких тысяч вычислений, проведенных последовательно для все более мелких решеток. По мере того как постоянная решетки уменьшается, силовые линии, «запертые» на грубой решетке (см. рисунок вверху), внезапно претерпевают фазовый переход и «рассыпаются» по всем направлениям. «Освобождение» силовых линий электрического поля служит хорошим доводом в пользу подхода, основанного на вычислениях с помощью решетки. Действительно, известно, что электрическое поле в обычном, непрерывном, пространстве-времени распространяется на бесконечно большие расстояния. Аналогичные вычисления, выполненные М. Кройтцем из Брукхейвена, показали, что с уменьшением постоянной решетки силовые линии хромоэлектрического поля по-прежнему остаются «запертыми», как в случае грубой решетки. Таким образом, удержание кварков может быть выведено из теории хромоэлектрического поля.

чтобы вероятность обнаружения одного из кварков протона была отлична от нуля только на трех узлах решетки. Когда постоянная решетки уменьшена вдвое по сравнению с ее первоначальным значением, распределение вероятностей должно быть уже сосредоточено на шести узлах. Уменьшение значения *g* снижает вероятность конфигурации полей на решетке, обладающих большим действием. Поэтому уменьшение *g* приводит к эффекту «крупного плана»: мы как бы стараемся рассмотреть с более близкого расстояния флуктуации поля, силовые линии и частицы на решетке. Таким образом, уменьшение значения *g* ведет к тому, что на решетке протон «выглядит» все большим.

Численное исследование свойств калибровочных полей проводят на решетке, занимающей большой, но все же ограниченный объем. Поэтому число узлов, граней и переменных, определяющих состояние, должно быть конечным, хотя оно может быть взято больше 100 000. Начальные значения всех переменных хранятся в памяти мощной ЭВМ. Изменяя по случайному закону элементы начальной конфигурации в соответствии со специальным алгоритмом, компьютер выборочно создает примерно 100 000 конфигураций. В результате вычисляется квантово-среднее эффектов, обусловленных конфигурациями из произведенной выборки.

Этот метод называют методом Монте-Карло. Еще до наших работ по удержанию кварков метод Монте-Карло был весьма успешно применен при анализе свойств термодинамических систем, и Вильсон подчеркивал, что его можно применять для изучения калибровочных теорий на решетке в квантовой механике. В 1979 г. в Брукхейвене мы с М. Кройтцем и Л. Джейкобсом впервые использовали метод Монте-Карло при исследовании абелевых калибровочных теорий. Мы хотели проверить, исчезает ли удержание частиц на решетке, наблюдаемое в КЭД при больших значениях константы связи, когда совершается переход к непрерывному пределу. Полученные результаты оказались очень интересными: на некоторой стадии уменьшения константы связи силовые линии на решетке внезапно претерпевали фазовый переход. При этом электрическое поле, «запертое» при больших значениях *g* между двумя электрическими зарядами на ребрах решетки, вдруг «развернулось» по всем направлениям вокруг зарядов.

Полезно сравнить неожиданное «освобождение» электрического поля со скачкообразным изменением свойств, наблюдаемым при переходе твердых тел в жидкую фазу. То, что Кройтц, Джейкобс и я обнаружили в численных расчетах, напоминало фазовый переход. Мы как будто имели дело с четырехмерным кристаллом, который могли охлаждать или нагревать, изменяя значение *g*. В некоторой точке наш кристалл претерпел фазовый переход; до перехода электрический заряд был «заперт», а после — «освободился». Впоследствии наши результаты были независимо подтверждены многими расчетами, выполненными в Европейской организации ядерных исследований (ЦЕРН) Б. Лаутрупом из Института Нильса Бора и М. Науенбергом из Калифорнийского университета в Санта-Барбаре, а также другими исследователями.

Вскоре после того, как мы получили свои результаты для абелевой калибровочной теории, Кройтц распространил метод Монте-Карло на неабелеву модель. Его результаты были столь же захватывающими, а с физической точки зрения даже более интересными. Правильный ответ для КЭД был известен и до наших работ — моделирование скорее проверяло метод Монте-Карло. Результаты же для неабелевых систем не были получены. Расчеты Кройтца показали, что в отличие от абелева случая в неабелевых калибровочных теориях на решетке при постепенном уменьшении константы связи фазовый переход не происходит. Наконец-то удалось показать, хотя и численными методами, что квантовая хромодинамика действительно ведет к удержанию кварков*.

С помощью метода Монте-Карло удалось исследовать предсказания квантовой хромодинамики для многих физических процессов. Например, после того как Кройтц показал отсутствие «освобождающего» фазового перехода в КХД, ему удалось оценить также силы, удерживающие кварки вместе; это хорошо согласуется с предсказаниями модели струны, которая не претендует на полное описание динамики кварков и адронов. Полученное Кройтцем значение для натяжения струны было независимо подтверждено Вильсоном, Ж. Бхано (ЦЕРН), мной и другими учеными.

Еще один удивительный результат, полученный методом Монте-Карло: при очень высоких температурах кварки перестают удерживаться вместе и могут двигаться почти свободно. Температура должна быть порядка двух триллионов градусов Цельсия — слишком высокая для получения в лабораторных условиях, но, возможно, не такая уж высокая, чтобы нельзя было найти свободные кварки в горячих звездах. По-видимому, свободные кварки могли существовать короткое время после так называемого Большого взрыва, послужившего началом современной фазы расширения Вселенной.

Сейчас передний край в исследованиях взаимодействий кварков — это численное моделирование изменений хромоэлектрического поля, происходящих при движении кварков. Во всех описанных исследованиях предполагалось, что кварки стационарны, и оценивалось хромоэлектрическое поле. Можно использовать метод Монте-Карло для моделирования квантовой механики кварков, движущихся в калибровочном поле. Однако объем необходимых вычислений делает такое моделирование пока невозможным даже для самых мощных компьютеров. Г. Хамбер из Принстонского института перспективных исследований, Э. Маринари и Дж. Паризи из Римского университета и я, а также независимо Л. Вейнгартен из Университета шт. Индиана недавно предложили приближенную схему для таких вычислений. Это приближение было использовано для расчета масс некоторых адронов; полученные теоретические результаты хорошо согласуются с экспериментальными значениями масс адронов.

Калибровочная теория поля на решетке подвела наконец квантовую хромодинамику к той ступени развития, когда появилась возможность делать на ее основе численные расчеты и сравнивать их результаты с экспериментальными данными. Это —элегантная теория, основанная на относительно простых математических представлениях, но богатая физическими следствиями. К настоящему времени все ее предсказания, которые можно было проверить, прошли испытание экспериментом. Теперь еще остается узнать, каким образом выводы квантовой хромодинамики могут быть доказаны логическими методами. В других областях теоретической физики есть много примеров, когда определенные результаты, полученные сначала численно, вскоре были доказаны аналитическими методами. Некоторые мои коллеги да и я сам уверены, что скоро квантовая хромодинамика так же успешно пройдет и это испытание.

---

* Следует отметить, что при уменьшении постоянной решетки константа связи ведет себя по-разному в случаях абелевой и неабелевой теорий. В неабелевой теории она уменьшается и стремится к нулю при уменьшении ячеек решетки. В этом проявляется асимптотическая свобода — свойство кварков двигаться почти свободно, находясь на достаточно близких расстояниях в адроне.

В квантовой электродинамике (абелевой калибровочной теории) асимптотической свободы нет. Взаимодействие между электроном и позитроном не ослабляется при их сближении. Результаты Кройтца показывают, что в этом случае при переходе к непрерывному пространству-времени константа связи стремится к некоторому постоянному значению. — *Прим. перев.*

Будущее сельского хозяйства США

В предстоящие десятилетия уровень цен на сельскохозяйственную продукцию в США будет определяться не только наличием таких основных ресурсов, как земля, вода и энергия, но и экспортным спросом

САНДРА С. БЕЙТИ И РОБЕРТ Г. ХИЛИ

СЕЛЬСКОЕ хозяйство в США уже в течение длительного времени является одной из самых процветающих отраслей экономики. Высокий уровень механизации в этой отрасли позволил значительно сократить долю занятой в ней рабочей силы и достичь большой производительности труда. Всего 3,1% рабочих производят со сравнительно низкой себестоимостью продукты питания и другую сельскохозяйственную продукцию в объемах, не только достаточных для удовлетворения внутренних потребностей, но и со значительными излишками, подчас вызывающими озабоченность. В последние три года экспорт сельскохозяйственной продукции из США составлял в денежном выражении более 40 млрд. долл. в год. В прошлом десятилетии эта цифра не превышала 10 млрд. долл. Валютные поступления от этого экспорта позволяли американской экономике сохранять относительную устойчивость в условиях различных внешних потрясений; начиная с 1973 г., к примеру, более 62% расходов страны на импорт нефти покрываются валютой, получаемой за проданную в другие страны продукцию сельского хозяйства.

Несмотря на такие внушительные достижения, существует все же озабоченность по поводу способности американского сельского хозяйства и в дальнейшем если не расширять, то хотя бы поддерживать достигнутый уровень производства и сохранять стабильные цены. Известно, что у США больше нет того резерва основных сельскохозяйственных ресурсов — земли, воды и энергии, — которым они когда-то располагали. Кроме того, одной из существующих проблем является снижение плодородия почв в ряде районов страны вследствие эрозии. Нет никакой гарантии, что в будущем вызванный этим обстоятельством спад продуктивности можно будет компенсировать повышением урожайности за счет улучшения генетики культур, внесением большего количества удобрений или другими агротехническими приемами, как это делалось раньше. Более того, быстрый рост цен на некоторые виды ресурсов, необходимые для сельского хозяйства, особенно на энергию, снизит рентабельность многих применяемых в настоящее время земледельческих приемов. Эти и другие факторы, такие, как широкое использование монокультуры (длительное выращивание культуры одного вида на одном и том же участке) или вероятность значительных климатических изменений, могут усилить уязвимость американского сельского хозяйства.

Как экономисты, проявляющие профессиональный интерес к этой проблеме, мы приняли участие в недавно проводившемся межотраслевом исследовании факторов, определяющих долгосрочную перспективу развития сельского хозяйства в США. Исследование проводилось по инициативе Федерального агентства по управлению в чрезвычайных обстоятельствах и Фонда охраны природных ресурсов. И хотя в оценке положения у экспертов не было единодушного мнения, многие из них проявили, как мы называем, сдержанный оптимизм. Что послужило основой для таких оценок, мы поясним далее, когда коснемся факторов, влияющих на производство продуктов питания в США, — количества и качества основных средств производства, используемых в сельском хозяйстве, темпов внедрения научно-технических достижений и преимуществ, связанных с наличием благоприятных климатических условий.

В ВЫСКАЗЫВАЕМЫХ в настоящее время сомнениях относительно нашей способности и в будущем производить продукты питания в достаточном объеме нет ничего нового. Вопрос об обеспечении продовольствием растущего населения мира не раз возникал в ходе истории. В 1798 г. Т. Мальтус выдвинул нашумевший тезис о том, что численность населения растет быстрее, чем средства к существованию. Соответствие между численностью населения и его обеспеченностью продовольствием, и то только на уровне, едва достаточном, чтобы не умереть с голоду, согласно Мальтусу, «должно регулироваться войнами, голодом, эпидемиями или различными формами «нравственной сдержанности»».

И хотя еще миллионы людей в мире не получают достаточного питания, этой мальтузианской ловушки удалось избежать. Ограничения в повышении сельскохозяйственного производства преодолены с помощью целого ряда агротехнических приемов: механизации, выведения улучшенных сортов растений и пород животных, увеличения доз вносимых удобрений, использования эффективных мер борьбы с сельскохозяйственными вредителями и болезнями, а также путем расширения сельскохозяйственных угодий за счет ирригации, обезлесения и дренажных работ. В результате менее производительные ресурсы стали использоваться более эффективно. Например, в Соединенных Штатах относительная нехватка рабочей силы сначала стимулировала использование наиболее плодородных земель, а позднее внедрение сельскохозяйственной техники. В начале прошлого столетия, еще задолго до того, как стала распространяться механизация сельского хозяйства, на производство примерно 3 т пшеницы требовалось 250 рабочих часов и два гектара земли. В настоящее время производство такого количества пшеницы требует 4 часов и одного гектара.

Из тех факторов, которые играют определяющую роль в дальнейшем развитии американского сельского хозяйства, первым мы рассмотрим площади пахотных земель. В последнее время уделяется большое внимание проблеме, связанной с использованием сельскохозяйственных угодий для иных целей, и в первую очередь для городского строительства. Согласно наиболее достоверным оценкам, сделанным М. Брюэром и Р. Боксли, в результате проведенного по инициативе федерального правительства обследования национального земельного фонда в период с 1967 по 1975 г. из сельского хозяйства ежегодно изымалось более 360 тыс. га пахотных зе-

мель под различные нужды, не связанные с производством сельскохозяйственной продукции. Это составляет 0,17% общей площади пахотных земель США, которая примерно равна 221 млн. га. Если площадь сельскохозяйственных угодий будет и дальше сокращаться такими же темпами, то к 2000 г. она уменьшится на 9,0 млн. га. Не слишком ли высоки эти темпы? Не ставится ли, таким образом, под угрозу обеспеченность страны продуктами питания в будущем? При рассмотрении этих вопросов следует иметь в виду, что, по данным министерства сельского хозяйства США, фактическая посевная площадь страны в 1981 г. составила 156,5 млн. га. Ранее наибольшая площадь посевных земель была в 1949 г. — 154,8 млн. га, наименьшая — в 1962 г. — 152,4 млн. га. Прогнозы будущих потребностей в посевных площадях обычно основаны на анализе наиболее неблагоприятной ситуации: они начинаются с оценки заниженных темпов повышения урожайности и более высоких темпов роста потребности в продовольствии. При этом исходят из допущения о неизменности цен на продукты питания по отношению к ценам на другую продукцию. Рост цен на сельскохозяйственную продукцию ограничил бы потребность в посевной площади, но не столько за счет снижения потребления продуктов в пересчете на калории, сколько в результате снижения потребления мяса; иными словами, значительная часть посевных площадей, используемых в настоящее время под выращивание кормов для крупного рогатого скота, свиней и птицы, могла бы использоваться более эффективно для производства продуктов питания.

Расчеты показывают, что потребность в дополнительной посевной площади на 2000 г. составит от 10,5 до 45 млн. га. Если допустить, что потенциальные посевные площади можно будет ввести в оборот без чрезмерно высоких экономических затрат или затрат на охрану окружающей среды и что некоторая часть пастбищ может быть переключена на выращивание сельскохозяйственных культур, то полученной в результате этих мероприятий дополнительной площади (60 млн. га) будет достаточно, чтобы удовлетворить потребность на конец нынешнего столетия, даже если площадь изъятой земли из сельского хозяйства к тому времени достигнет 9 млн. га. Этой дополнительной площади будет достаточно даже в том случае, если при расчете потребности в земле учесть высокие экспортные поставки, сравнительно низкую урожайность и неизменность относительных цен на сельскохозяйственную продукцию.

При проведении анализа не учитывалось, что значительная часть земель, которая рассматривается как резерв, уже используется под лесопосадки или под пастбища. Использование их в иных целях может привести к значительному сокращению производства древесины и мяса. Возможные побочные следствия таких мер почти не исследовались.

Более того, обобщенные статистические данные о посевной площади в США могут ввести в заблуждение, так как они не отражают существенных изменений в использовании земли в отдельных районах. В период с 1939 по 1978 г. в северо-восточных и юго-восточных штатах из сельского хозяйства была изъята примерно треть земель, в то время как в других районах страны наблюдалось увеличение площадей под выращивание сельскохозяйственных культур. На востоке страны значительная часть посевных площадей была изъята не для строительства жилья, торговых центров и т.п., а для лесопосадок. Как отмечает Ф. Рауп, экономист в области сельского хозяйства и эксперт по вопросам землепользования из Миннесотского университета, «главным претендентом на сельскохозяйственные угодья является промышленность, которая наносит ущерб не столько изъятием земель, сколько отвлечением из сельского хозяйства рабочей силы. Урбанизация создает потребность в земле на создание зон отдыха, строительство жилых домов, а это неразрывно связано со стихийным расширением частных лесных массивов, не используемых в производственных целях. Наши наиболее индустриально развитые районы и города стали самыми лесистыми зонами. В конкуренции за сельскохозяйственные земли проявляется очевидный факт: предпочтение отдается не производству продуктов питания на местах, а выращиванию деревьев». Кроме того, на юго-востоке США часть пахотных угодий превращена в пастбища, поскольку Юг стал важным районом выращивания крупного рогатого скота для откормочных хозяйств Среднего Запада.

Другим крупным претендентом на пахотные угодья являются магистральные шоссейные дороги, строительство которых осуществляется в соответствии с Федеральной программой. Значительная часть дорожных магистралей, связывающих многие штаты, построена за двадцатилетний период с 1956 по 1975 г. Трудно переоценить то огромное влияние, которое эта гигантская программа оказала на систему землепользования в стране. Под новые шоссейные трассы был использован почти 1 млн. га земли, и помимо этого появилась возможность городского и промышленного строительства еще на миллионах гектаров в районах, прилегающих к этим дорогам. Большая часть данных, иллюстрирующих изменения в системе землепользования США, была собрана в тот период, когда последствия дорожно-строительного бума ощущались наиболее сильно. Однако эти данные не могут служить надежной основой для прогноза масштабов изъятия земель в будущем, ибо Соединенные Штаты вряд ли предпримут новую грандиозную программу расширения своей дорожной сети.

Почти одновременно со строительством дорожных автомагистралей проводилась в жизнь и крупномасштабная программа строительства



ОТНОСИТЕЛЬНЫЕ ТЕМПЫ РОСТА продуктивности сельского хозяйства США за последние два десятилетия достигли исключительно высокого уровня. Показатель общей продуктивности рассчитан с учетом основных производственных ресурсов и объемов производства сельскохозяйственной продукции. (Данные министерства сельского хозяйства США.)

плотин. Огромные площади пахотных земель оказались под водой, а многие примыкающие к водохранилищам районы стали непригодными для использования в сельском хозяйстве. Например, в бассейне реки Миссури расположенные в низинах луга, которые сейчас затоплены, использовались для выращивания кормов на зимний период, а скот пасли летом на более возвышенных сухих местах. Сейчас эра строительства плотин подходит к концу.

В отличие от дорожного и гидротехнического строительства влияние урбанизации, по-видимому, еще не проявилось в полной степени. Пока оно только ощущается в стирании различий между традиционными способами землепользования. По мере того как сельские районы вновь заселяются мигрантами из городов и пригородов, происходит стирание различий между фермами и жилыми домами в сельской местности, и возникают новые виды смешанного землепользования. Большинство сельских жителей уже не являются фермерами: на фермах проживает всего 10% сельского населения. Разность интересов, характерная для такого разнородного общества, еще более способствует снижению рентабельности сельского хозяйства и дополнительному изъятию земель из сельскохозяйственного оборота.

ОЖИДАЕТСЯ, что потребность в жилищном строительстве в сельских районах будет расти. С 1980 по 1990 г. примерно 40 млн. американцев, родившихся в период «бума младенцев» (50-60-е годы), достигнут двадцатипятилетнего возраста, и в этой связи спрос на жилье может возрасти до такого уровня, какого в США не наблюдалось ни в одно из предыдущих десятилетий. Одновременно ожидается увеличение потребности и в земельной площади под строительство заведений для отдыха и развлечений и домов для престарелых. Это обстоятельство еще более усугубляет конкурентную борьбу за потенциальные сельскохозяйственные площади. Выделение земель на эти цели в ущерб сельскому хозяйству — это лишь одна сторона рассматриваемого процесса; имеются и другие, хотя и не столь очевидные негативные явления, как следствие того, что земельные угодья дробятся на отдельные экономически нерентабельные участки, а в их использовании нет единого подхода. Все это сдерживает эффективную эксплуатацию разбросанных участков фермами, к которым они примыкают.

Серьезным претендентом на сельскохозяйственные земли становится энергетика. Эта отрасль использует земельные площади для различных целей — от добычи угля до получения продуктов топлива из растений и растительных отходов. Растут масштабы открытой разработки угля, особенно в штатах на территории Великих Равнин, где угольные пласты залегают на глубине 30 м. Расчеты показывают, что в США до конца столетия добыча угля открытым способом будет ежегодно поглощать 227 тыс. га. Конечно, значительная часть этой площади намечена к восстановлению. И хотя восстановление земель потребует времени и значительных материальных затрат, а земли при этом не обретут прежнего плодородия, потери в производстве продуктов питания по этой причине не должны быть существенными. Помимо этого, существует опасение, что разработка шахт, строительство электростанций и работы, связанные с производством синтетического топлива, потребуют большого количества воды из бассейнов Великих Равнин, что дополнительно ограничит производство продовольствия в этом регионе.

Более серьезным фактором, который повлияет на производство продуктов питания в стране, является планируемое выращивание культур для получения топлива, например газохола, в состав которого входит спирт. Представляется, однако, что спирт, получаемый из зерновых, не будет конкурировать с другими видами топлива до тех пор, пока на его производство не будут выделяться дополнительные средства. Добавка 10% спирта к общему объему ежегодно сжигаемого в США бензина (380 млрд. л) означает, что на производство требуемого количества спирта будет уходить 60% урожая зерновых в стране. Очевидные последствия этого мероприятия должны охладить наш пыл. Администрация Рейгана пока не проявляет стремления выделять субсидии на переработку биомассы, и этот заменитель топлива в ближайшем будущем не будет являться угрозой для сельского хозяйства.



ПЛОЩАДЬ ПАХОТНЫХ ЗЕМЕЛЬ, фактически используемых в сельскохозяйственном производстве США, была рекордной в 1981 г. и составляла 156,5 га. Резкое увеличение этого показателя произошло за последнее десятилетие. Площадь земель, которые дополнительно могут быть введены в оборот, включая и те, что в настоящее время используются под пастбища, составляет около 60 млн. га. Таким образом, потенциальная площадь обрабатываемых земель равна примерно 221 млн. га. (Оценка сделана сотрудниками министерства сельского хозяйства США.)

КРОМЕ земельных угодий, имеются еще два вида важных ресурсов, от которых зависит производство продовольственных продуктов, — вода и энергия. Для повышения урожайности сельскохозяйственных культур во многих западных районах США существенное значение имеет ирригация. Орошение способствует повышению урожайности, позволяет получать двойные урожаи и дает возможность фермерам выращивать более ценные культуры. По сравнению с 1940 г. площади орошаемых земель в целом по стране утроились, и в стоимостном выражении на них производится почти четверть всего объема сельскохозяйственных культур. Значительное увеличение орошаемых площадей объясняется тем, что во многих районах на протяжении долгого времени считалось, что вода — это товар, который не стоит ни цента, и любой фермер может иметь воду бесплатно, заплатив лишь за ее доставку к своему полю. А за счет низких расценок на все виды энергии и благодаря федеральным субсидиям, получаемым всеми фермерами, транспортировка воды стоила очень дешево.

Но эра дешевой воды отошла в про-

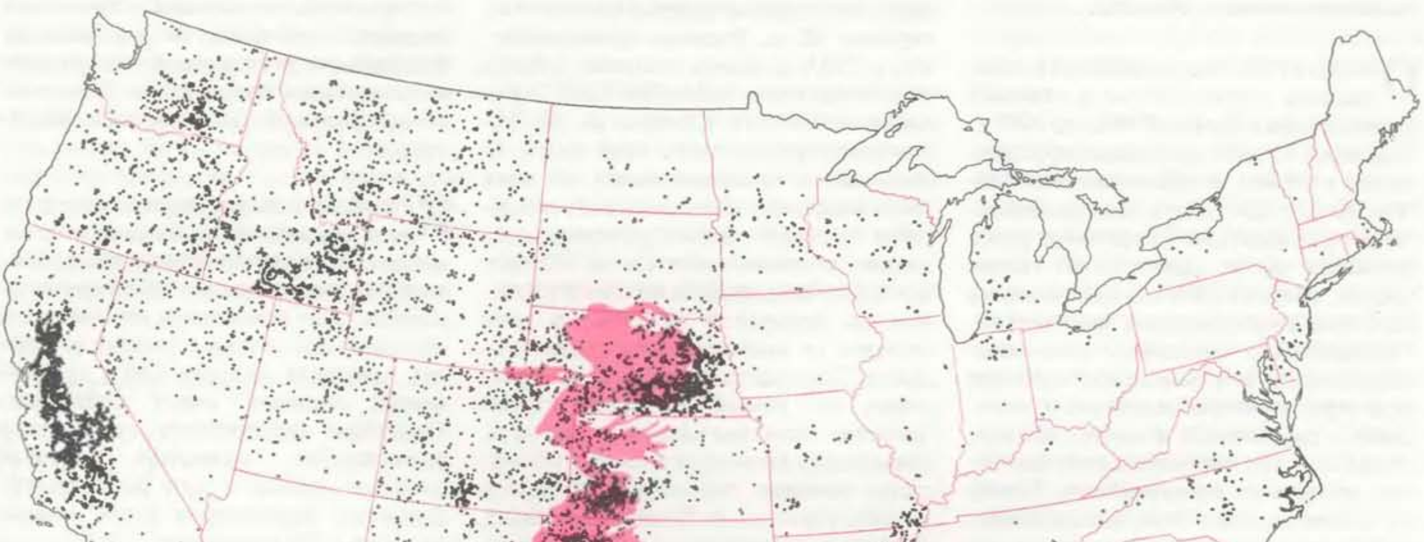
шлое. В тех районах, где фермеры вынуждены использовать грунтовые воды, откачка обходится все дороже, поскольку их уровень постоянно снижается. Подача воды из поверхностных источников тоже не всегда рентабельна, так как из-за высокой стоимости энергии доставка воды на большие расстояния теперь связана с колоссальными затратами. На западе страны большая часть запасов поверхностных вод уже распределена, и новым потребителям приходится все чаще покупать соответствующие права у теперешних хозяев.

Увеличение орошаемых площадей в последнее время имело место в четырех штатах: Небраске, Канзасе, Оклахоме и Техасе. Орошаемые земли используются здесь в основном для выращивания сорго, кукурузы и люцерны. Урожаи этих культур идут главным образом на корм скоту, так как в этих штатах находятся крупнейшие откормочные хозяйства страны. Производство говядины стало возможным здесь за счет разработки водоносного горизонта Огаллала — огромного подземного озера, простирающегося под территорией восьми штатов (*см. иллюстрацию, приведенную ниже*). Пополнение этого источника происходит чрезвычайно медленно, его уровень неуклонно падает, а стоимость откачки воды из него обходится все дороже.

В конечном счете из-за падения уровня грунтовых вод и роста цен на энергию уменьшатся масштабы орошения земель грунтовыми водами. Фермеры, пытаясь снизить расходы на воду, прибегают к почвозащитным мерам, поэтому рентабельность земледелия на орошаемых и богарных (без полива) землях почти одинакова. Если в эти районы вернется богарное земледелие, то в штатах, под территорией которых находится водоносный горизонт Огаллала, крупномасштабный откорм скота утратит свои преимущества и станет намного менее выгодным. Как указывал Рауп, «не имеющая никакой цены вода способствовала созданию такого уровня потребления говядины в пределах всей страны, поддерживать который долго вряд ли будет возможно, если только нам не придется вновь обратиться за помощью к фермерам Кукурузного Пояса, выращивающим кормовые культуры. Наша экономика сейчас зависит от уровня мясного скотоводства, а это значит, от истощаемых источников. И зависимость эта приобрела угрожающие размеры».

Независимо от того, что предстоит пережить нашему откормочному хозяйству, ясно одно: рост цен на воду должен быть компенсирован повышением цен на зерно, в противном случае ирригация не сможет беспредельно способствовать расширению сельскохозяйственного производства, как это имело место в последние три-четыре десятилетия. К. Фредерик, сотрудник Управления по ресурсам будущего, подчеркивает, что это явление — необязательно только негативное. Он пишет: «Если учесть, что в условиях дефицита воды повышается эффективность ее использования, то за счет ирригации можно будет наращивать сельскохозяйственное производство еще в течение многих десятилетий. Ситуация станет по-настоящему серьезной, если мы будем и впредь считать воду дешевой, хотя на самом деле она стоит дорого. Такая политика наверняка приведет к неэффективному использованию воды . . . и ограничит роль ирригации в дальнейшем развитии сельского хозяйства».

ЭНЕРГИЯ на фермах используется для различных целей. Помимо орошения, о котором мы уже говорили, сельское хозяйство ощущает еще боль-



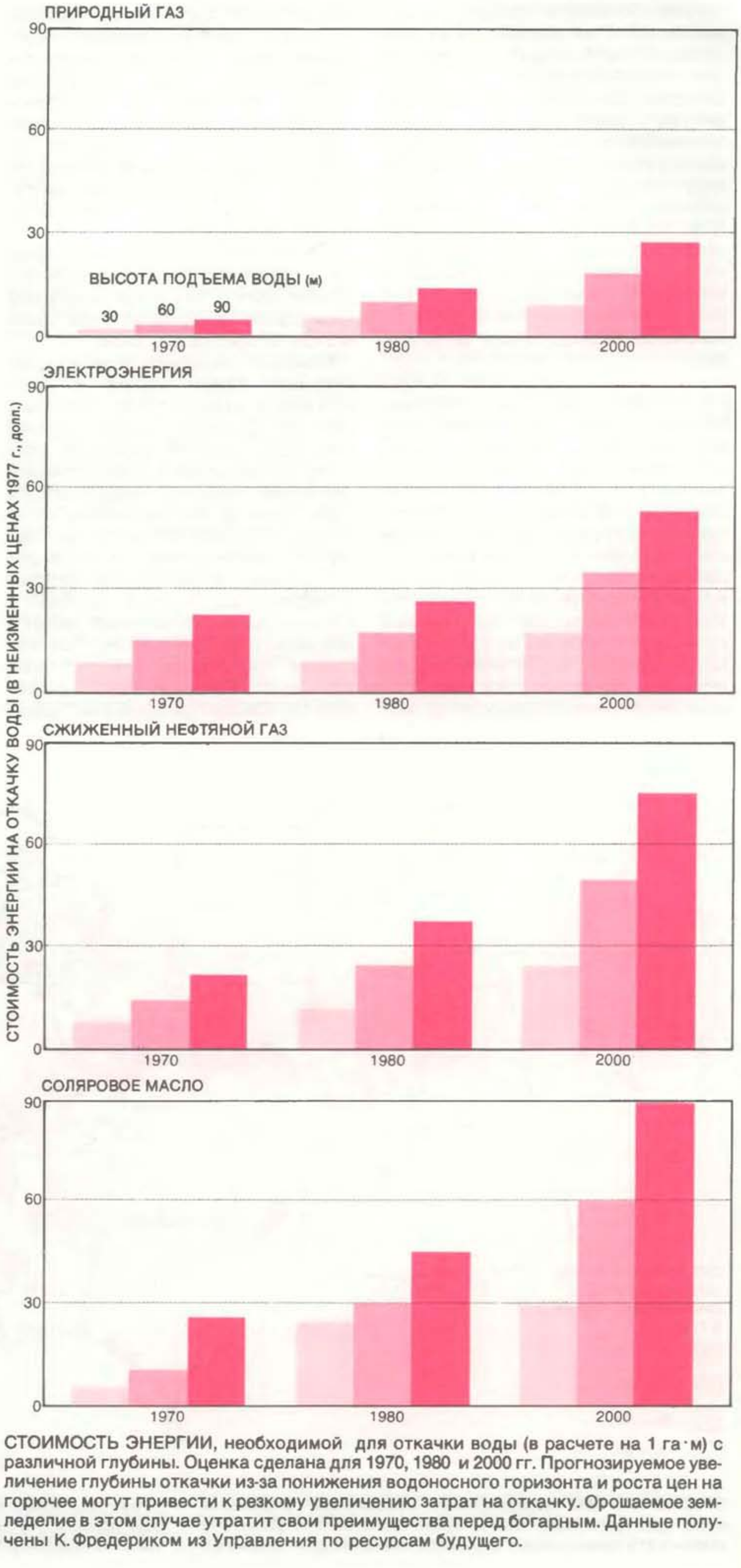
ОРОШАЕМЫЕ ЗЕМЛИ В США занимают площадь примерно 24 млн. га. Каждая точка на карте соответствует площади 3260 га, на которой действовали ирригационные сооружения в 1977 г. (последний год, на который имеется информация). Цветным участком отмечены земли, под которыми простирается водоносный горизонт Огаллала. Именно в этой зоне в основном происходило расширение орошаемых площадей в послевоенный период. Здесь главным образом выращиваются кормовые культуры.

шую потребность в различных видах энергии для внесения химических удобрений, а также в различных видах топлива. Потребление энергии в сельском хозяйстве постоянно росло. Гербициды сейчас применяются более чем на 90% площадей, занимаемых под кукурузой, а десятилетия назад гербицидами обрабатывалось 57% площадей, занятых этой культурой. Такая же картина наблюдается и в отношении механизации: в пересчете на л.с. парк тракторов по сравнению с 1966 г. к 1978 г. возрос на 31%.

Экономические причины такой тенденции легко понять. В период между 1950 и 1978 гг. повышение цен на энергию отставало от роста стоимости рабочей силы в сельском хозяйстве в два раза, а от роста цен на землю — в четыре раза. Поэтому фермеры стремились возместить расходы и на землю и на рабочую силу применением ядохимикатов, топлива и удобрений. Данные соотношения цен подкреплялись правительственной политикой поддержания цен на топливо и удобрения на возможно низком уровне. Цены на природный газ, например, удерживаются правительством на искусственно заниженном уровне, а электроэнергией большинство фермерских хозяйств обеспечивается на самых благоприятных условиях.

Дешевая энергия используется на фермах не только для уменьшения расходов на рабочую силу, но также для повышения отдачи всех прочих средств производства и для того, чтобы насколько возможно уменьшить потери урожая и улучшить плодородие полей. Этого можно достичь разными способами: применяя пестициды, можно ограничить колебания урожаев, ирригацией компенсировать неблагоприятное воздействие климата, а используя просушку семян или выдерживание их на холоде, можно обеспечить надежные всходы на полях.

Несмотря на постоянное увеличение стоимости энергии, ее потребление в сельском хозяйстве вряд ли будет снижаться, так как ее роль в этой сфере исключительно важна и многозначна. Нельзя сказать, что относительные цены на энергию растут такими темпами, которые нарушают нормальный ход производства продовольствия или оказывают существенное влияние на их цены. Но все же в связи с ростом стоимости энергии могут произойти некоторые изменения в территориальной структуре фермерских хозяйств. Более высокие транспортные расходы могут вызвать перемещение откорма скота преимущественно в южные районы страны, чтобы исключить перевозку животных в районы Среднего Запада. По той же причине выращивание различных культур на местах, например помидоров, может оказаться бо-



СТОИМОСТЬ ЭНЕРГИИ, необходимой для откачки воды (в расчете на 1 га·м) с различной глубины. Оценка сделана для 1970, 1980 и 2000 гг. Прогнозируемое увеличение глубины откачки из-за понижения водоносного горизонта и роста цен на горючее могут привести к резкому увеличению затрат на откачку. Орошаемое земледелие в этом случае утратит свои преимущества перед богарным. Данные получены К. Фредериком из Управления по ресурсам будущего.

лее рентабельным по сравнению с производством в тех районах, где их себестоимость хотя и ниже, но транспортировка обходится слишком дорого. Большую роль в повышении экономии энергии могут сыграть научно-технические достижения и меры административного характера; примерами могут служить применение более рациональных графиков подачи воды на поля, комплексный подход в планировании мер борьбы с сельскохозяйственными вредителями или переход на замкнутый технологический цикл с вторичной переработкой отходов на некоторых фермах, скажем на молочных.

О БОЛЬШИХ масштабах американского сельскохозяйственного производства могут свидетельствовать следующие цифры: из мирового урожая на долю США приходится 13% пшеницы, 29% кормового зерна (в основном кукурузы), 17% мяса и 62% сои. Поскольку в США проживает 5% населения мира, то даже при существующем высоком уровне внутреннего потребления продуктов собственного производства в стране остаются излишки на экспорт. Тем не менее следует отметить, что дальнейший рост сельскохозяйственного производства возможен только при условии сохранения плодородия земель. Главными проблемами являются борьба с ветровой и водной эрозией почвы, уплотнением почвы тяжеловесными сельскохозяйственными машинами, возрастающей засоленностью орошаемых земель, а также некоторыми побочными последствиями использования растительных продуктов для производства энергетических ресурсов.

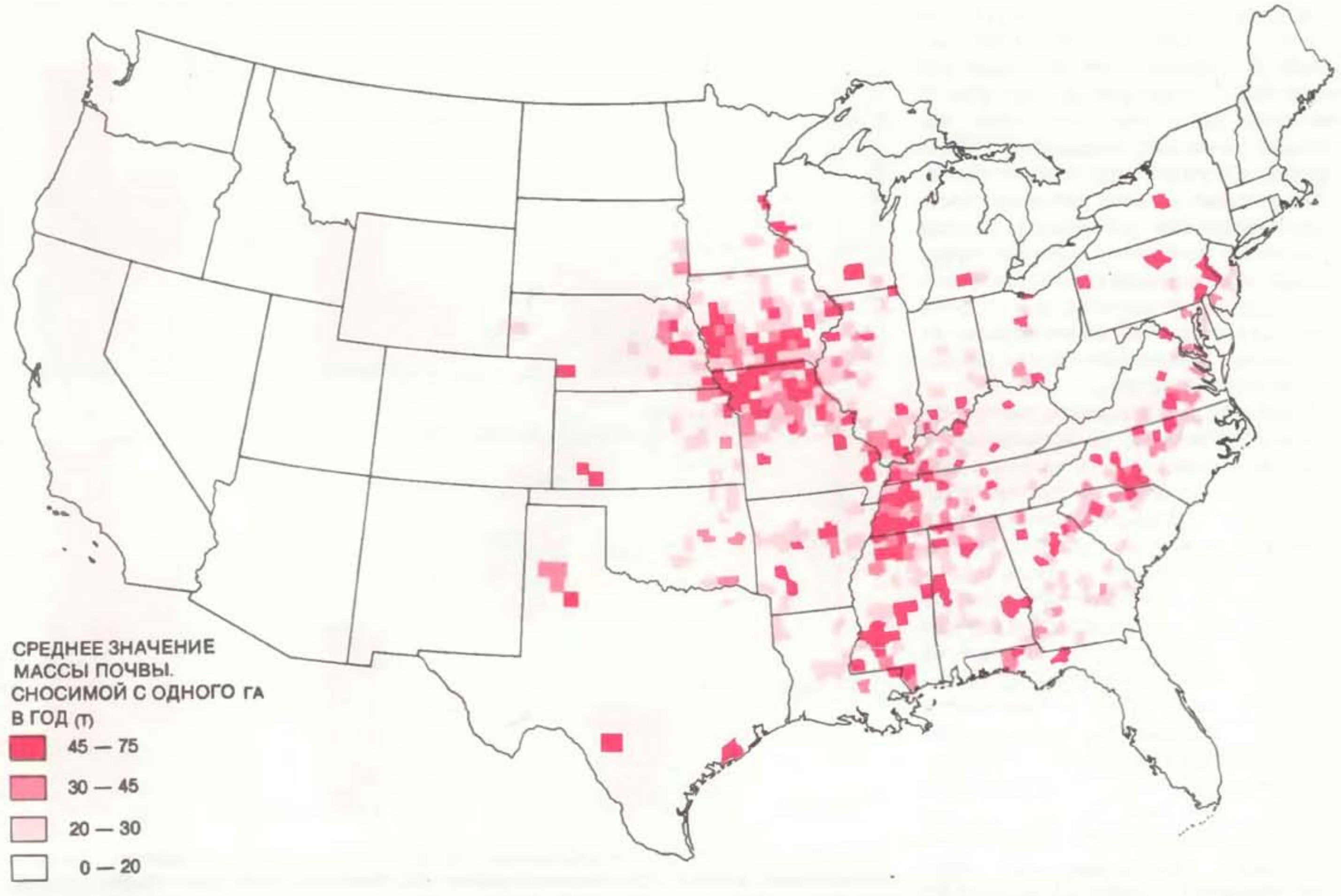
Эрозия может снести значительную часть верхнего слоя. При сносе почвы в объеме 100 т с 1 га в год почвенный слой снижается почти на 1 см. Такая потеря не заметна для глаза.

Наиболее серьезную угрозу представляет водная эрозия в юго-восточных штатах страны, где после каждого проливного дождя толщина пахотного слоя на убранном поле может снизиться на 2,5 см. Интенсивная эрозия, к счастью, наблюдается в США лишь на незначительной части пахотных угодий. Водная эрозия, связанная с производством сои, отмечается, например, в ряде крупных сельскохозяйственных областей на Юге и Среднем Западе. От ветровой эрозии в основном страдают районы Высоких Равнин в шт. Техас, юго-восточные области шт. Колорадо и юго-западная зона шт. Канзас, где преимущественно выращивают пшеницу (см. приведенные карты).

Превышает ли нынешняя интенсивность эрозии ту, что имела место в 30-е годы? Точно ответить на этот вопрос пока нельзя. Есть, правда, некоторые доказательства того, что эрозия в последние годы усилилась.

Д. Тиммонс из Университета шт. Айова указывает, что ежегодные потери почвы в штате уменьшились с 64 т/га в 1949 г. до 35 т/га в 1957 г., но в 1974 г. масштабы эрозии вновь возросли и достигли 42,6 т/га. Эта обратная тенденция в значительной мере объясняется тем, что в результате быстрого роста цен на пшеницу, кукурузу и сою в начале 70-х годов фермеры стали распахивать подверженные эрозии земли, ранее использовавшиеся под пастбища. Этот процесс стал еще более интенсивным после 1974 г.

Увеличение площадей под кукурузу, сою, пшеницу и хлопок в связи с ожидаемым повышением спроса на продовольствие и другую сельскохозяйственную продукцию может способствовать усилению эрозии, но ее можно ослабить, если принять соответствующие меры по защите почвы. Под этими мерами можно понимать любой способ обработки почвы, который уменьшает ее вымывание водой по



ЭРОЗИЯ ПОЧВЫ представляет наибольшую опасность лишь для небольшой части пахотных земель страны, но именно эти земли имеют огромное значение для производства продовольствия. Карты на этих двух страницах содержат данные о средней массе почв, снесенных в 1977 г. с одного гектара водной и ветровой эрозией. На карте слева по-

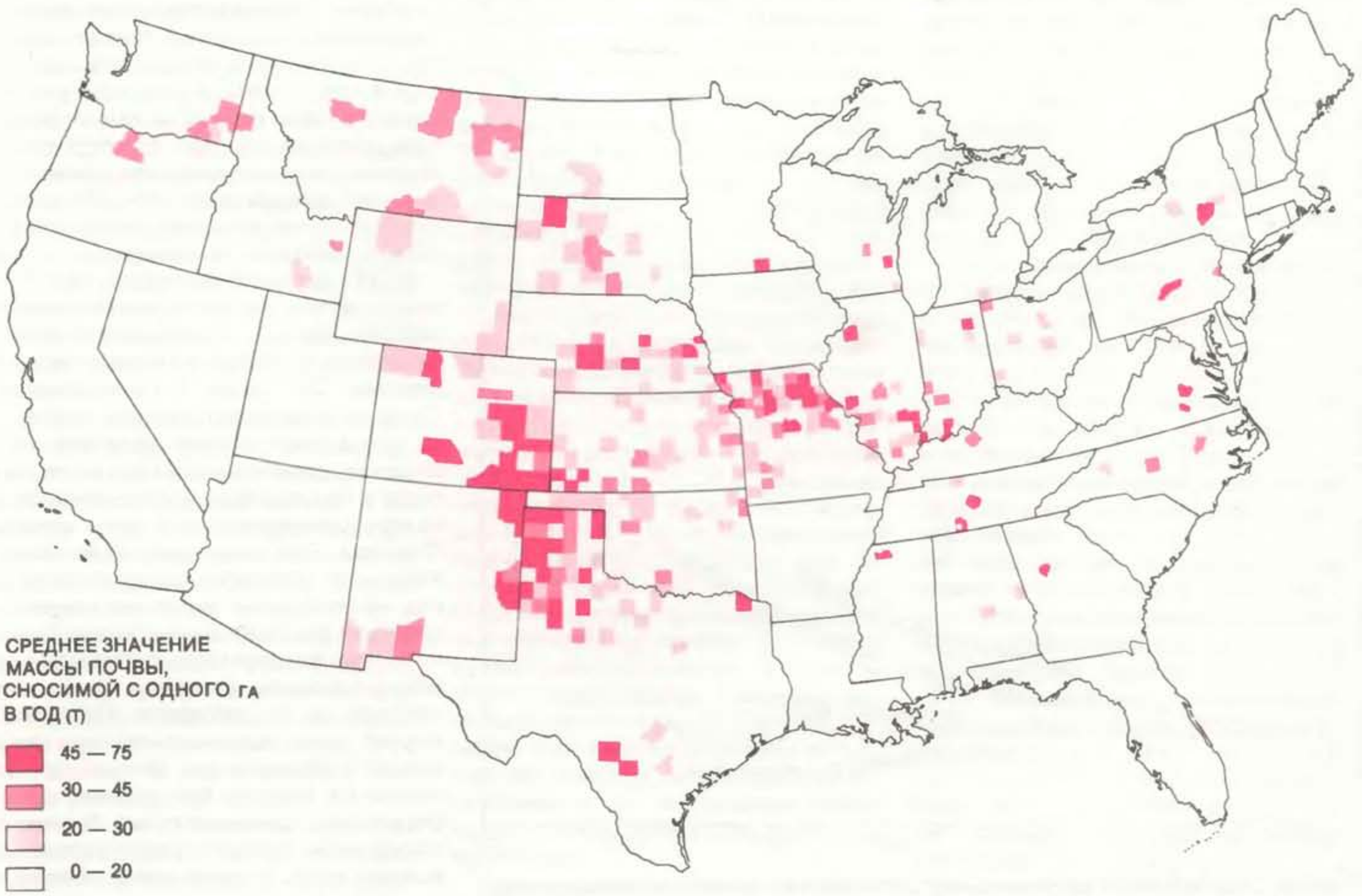
сравнению с традиционным выравниванием зяби. Обычные методы почвозащитной обработки включают нулевую и минимальную вспашку, гребневую посадку, вспашку-посев (при этом обработка почвы и посев семян производятся одним агрегатом), а также чизелевание. Начиная с 60-х годов распространение почвозащитной обработки происходило быстрыми темпами, и в настоящее время ее применяют на 25% всех обрабатываемых площадей. Активное распространение почвозащитных мер в основном объясняется экономическими причинами: они требуют меньших затрат на рабочую силу и горючее по сравнению с традиционными методами возделывания земли. К тому же результат налицо — эрозия снижается.

П. Кроссон из Управления по ресурсам будущего полагает, что к 2010 г. почвозащитной обработкой будет охвачено 50—60% пахотных земель страны. Из его подсчетов следует, однако, что площадь возделываемых пахотных земель может увеличиться к 2010 г. на 30 млн. га и «значительная часть этого прироста будет включать земли, подверженные более сильной эрозии по сравнению с теми, что используются для производства культур в настоящее время». В таком случае подвергаемые эрозии почвы будут преобладать и в дальнейшем.

ПОЧВЕННАЯ эрозия снижает плодородие земель, так как она уменьшает способность почвы задерживать влагу, уносит из нее питательные вещества для растений, ухудшает структуру и нарушает однородность почвы в пределах одного поля. Эти негативные явления не всегда одинаковы и зависят от условий в данном регионе, характера почвы, применяемой агротехники, вида возделываемой культуры и продолжительности вегетационного периода. Оценить ущерб, который эрозия причиняет урожаю, — дело не легкое, поскольку он зависит от многих других факторов. По мере того как почва с годами все больше истощается эрозией, фермеры увеличивают дозы удобрений, чтобы как-то компенсировать потери пахотного слоя и повысить плодородие земель. Они применяют иные агротехнические приемы, используют другие виды техники и меняют виды культур.

Несмотря на эти меры, усложняющие ведение хозяйства, эксперты, изучавшие влияние эрозии на продуктивность земель, пришли к выводу, что в ряде случаев потеря пахотного слоя даже на толщину в 2,5 см уже может оказать существенное влияние на урожайность. Там, где эрозия намного уменьшила пахотный слой, корневая система растений развивается в нижних слоях почвы, менее богатых питательными веществами. Расходы на выращивание культур на таких землях повышаются, так как фермеры вынуждены вносить больше удобрений для поддержания высокой урожайности. Добавка же питательных веществ в виде удобрений может повысить урожайность только в том случае, если в почве имеется достаточно влаги. Таким образом, различия в способности почв удерживать необходимую растениям влагу в основном определяют различия в урожайности.

Снижая производительность земель, эрозия почв к тому же в большой степени способствует загрязнению воды. Сток с полей содержит нерастворимый осадок, элементы минерального удобрения, гербициды и пестициды — все они являются загрязнителями, если содержатся в воде в чрезмерно больших количествах. Нерастворимые осадки, смываемые с фермерских полей, сокращают продолжительность жизни многих озер и водохранилищ и увеличивают затраты на поддержание русл судоходных рек. Большое количество элементов минеральных



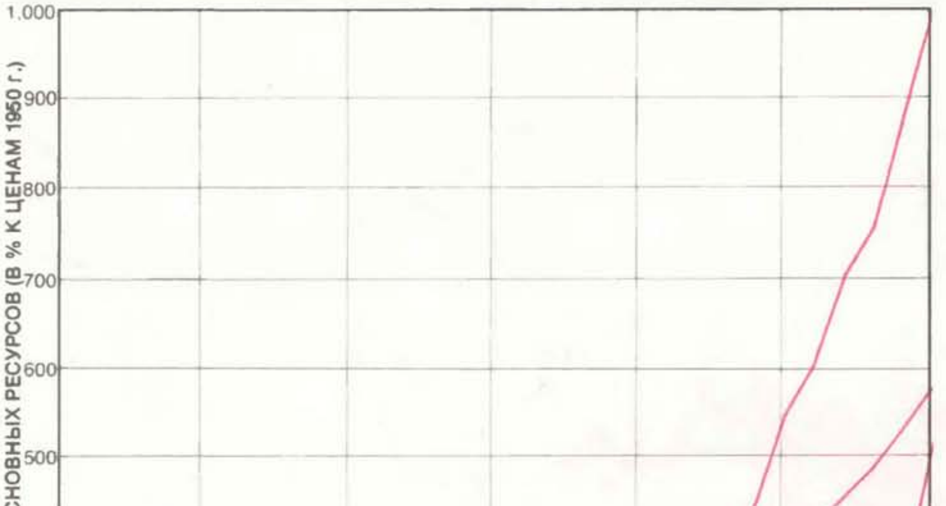
казаны масштабы эрозии на землях, используемых для выращивания сои; в основном это водная эрозия. Карта справа отражает масштабы преимущественно ветровой эрозии на землях, используемых для посева пшеницы. Условные знаки обозначают интенсивность эрозии.

удобрений повышает допустимую концентрацию вредных примесей в питьевой воде и затраты на ее очистку. Гербициды и пестициды наносят ущерб живой природе. Но поскольку эти вредоносные явления имеют место за пределами фермерских хозяйств, то их хозяева не утруждают себя заботой о последствиях.

СУЩЕСТВУЮТ и технические решения проблемы почвенной эрозии. К ним относится устройство террас на крутых склонах, оставление в поле растительных остатков после сбора урожая, применение почвозащитной обработки земли и контурной вспашки. Для стоков с полей можно устраивать водосборы. Результаты большинства исследований показывают, что почвозащитная обработка снижает эрозию на 50 — 90% в зависимости от того, какая часть поля осталась покрытой растительными остатками.

Если эти меры борьбы с эрозией столь эффективны, то почему они не получили еще более широкого распространения? Причины разные — это и субъективное отношение к ним фермеров, и физические свойства почвы, и характер землевладения, и политика налогообложения, и программы производства сельскохозяйственных продуктов. Главная же причина кроется просто в том, что многие почвозащитные приемы экономически не оправдываются. Учитывая, что эрозия снижает плодородие земель и загрязняет воду, а фермеры проявляют слабую заинтересованность в применении почвозащитных мер, а также принимая во внимание, что с увеличением экспортных поставок продовольствия эксплуатация земель становится более интенсивной и в связи с этим эрозия может возрастать, многие обозреватели пришли к выводу, что проблема почвенной эрозии заслуживает более серьезного внимания со стороны общественности и требует действенных мер по борьбе с эрозией.

ПЛОДОРОДИЕ земли снижается и в результате применения тяжеловесной сельскохозяйственной техники, которая вызывает уплотнение почвы и, как следствие, ухудшение просачивания влаги, что ведет к осушению почвы, быстрому стоку воды с поверхности, а также препятствует прорастанию корней. Некоторые культуры, например картофель и сахарная свекла, более чувствительны к этим явлениям, чем другие культуры. Уплотнение почвы ухудшает ее аэрацию, затрудняет передачу тепла, что препятствует росту большинства растений.



ОТНОСИТЕЛЬНЫЙ РОСТ ЦЕН за период 1950 — 1981 гг. на четыре основных вида ресурсов, используемых в сельском хозяйстве, — землю, энергию, рабочую силу и удобрения. Наиболее быстрыми темпами росли цены на землю; за последние тридцать лет они увеличились более чем на 900%. (Данные министерства сельского хозяйства США.)

Известны случаи падения урожайности в результате уплотнения почв, но имеются и другие данные, свидетельствующие об обратном. Например, урожаи сои в шт. Миннесота на умеренно уплотненных почвах бывают выше, чем на неуплотненных. В действительности мы очень мало знаем о том, какие площади подвержены уплотнению и как оно влияет на урожай. Но, несмотря на это, имеются агротехнические методы, позволяющие предупредить нежелательные последствия, к которым может привести это явление. К ним относятся строгое соблюдение сроков сельскохозяйственных работ, сокращение выездов в поле, соблюдение движения колесного транспорта по проторенной колее, проведение более глубокой вспашки и применение широкозахватной техники. Результаты обследования, проведенного недавно управлением конгресса США по техническим оценкам, показали, что существующая технология вполне способна в значительной мере снизить влияние фактора уплотнения почвы на ее плодородие.

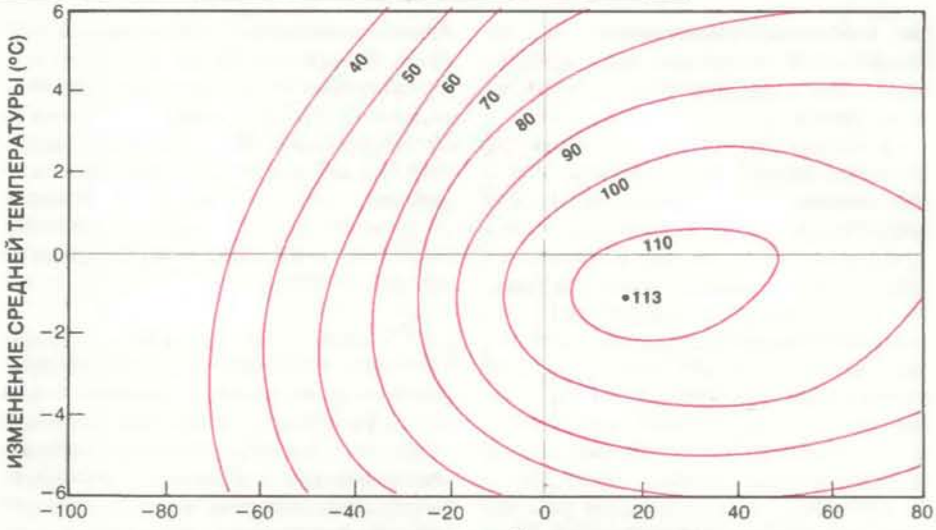
Повышенная засоленность почвы в сфере охраны окружающей среды рассматривается как наиболее актуальная проблема, порождаемая практикой орошаемого земледелия. Чистая вода после полива испаряется и просачивается в глубь почвы, а соль остается в поверхностном слое. Если она не смоется дождями или при последующих поливах, то это отрицательно скажется на плодородии земли. 25—30% орошаемых земель на западе страны считаются чрезмерно засоленными.

В США имеются два района, в которых проблема засоления почвы стоит наиболее остро — это бассейн нижнего течения Колорадо и западная часть равнины Сан-Хоакин в Калифорнии. Согласно имеющимся оценкам, в районе Сан-Хоакин урожаи снизились на площади, превышающей 160 тыс. га, в связи с чем ежегодные экономические потери составляют 31,2 млн. долл. Фермеры обратились с просьбой к правительству выделить капиталовложения на проведение работ по улучшению качества воды, но до решения этого вопроса фермеры могут сами исправить создавшееся положение другими, доступными им способами. Они, например, могут применить системы капельного орошения или прорыть оросительные каналы, что повысит эффективность орошения полей. К тому же они могут улучшить систему рециркуляции воды. С этой целью вначале воду следует использовать для полива тех полей, на которых выращиваются растения, наиболее чувствительные к соли. В последующем той же водой (по

мере увеличения содержания соли в ней) поливают поля с более солевыносливыми культурами. Несмотря на то что вопрос о выделении средств для борьбы с засолением почвы еще не решен, финансировать эти мероприятия, хотя бы частично, необходимо уже сейчас.

ПОВЫШЕНИЕ цен на энергетические ресурсы вынудило многих осуществить поиск недорогих способов получения энергии. Один из них — использование растительных остатков непосредственно в качестве топлива или в качестве сырья для получения топлива. И хотя эта идея технически осуществима и растительные остатки имеются в огромных количествах (по некоторым оценкам, 400 млн. т в год в сухом весе), на практике она может оказаться экономически нецелесообразной или нежелательной. Растительные остатки неудобны для транспортировки, они объемны, и доставка их к месту переработки может обойтись недешево. Расчеты показали, что использование растительных остатков различных сельскохозяйственных культур в качестве топлива обойдется не дешевле, чем угля, если принять во внимание стоимость их перевозки к месту потребления. Вместе с тем, если растительные остатки оставить в поле, они уменьшат эрозию, а содержащиеся в них питательные вещества будут вновь внесены в почву. Поэтому во многих районах фермеры предпочитают не убирать их с полей. Короче говоря, этот вид сельскохозяйственной продукции нельзя рассматривать как потенциальный энергетический ресурс ни с точки зрения экономики, ни с точки зрения охраны окружающей среды. Поэтому упомянутые 400 млн. т никогда полностью не идут на переработку. По имеющимся оценкам, лишь 80 млн. т могут быть использованы на эти цели.

И хотя считается признанным, что состояние американского сельского хозяйства в перспективе в основном будет определяться количеством и качеством земли и других природных ресурсов, мальтузианского призрака пока все же удавалось избегать главным образом благодаря научно-техническим достижениям, внедрение которых способствовало существенному повышению продуктивности при данных ресурсах. При этом в настоящее время нет полной уверенности, что нам удастся и далее поддерживать темпы роста продуктивности сельского хозяйства на уровне нескольких прошлых десятилетий. Даже некоторые наиболее оптимистически настроенные исследователи полагают, что если Соединенные Штаты смогут в будущем сохранить прежние темпы роста только урожайности в абсолютном выражении, то это уже будет хорошо.



ВЛИЯНИЕ КЛИМАТИЧЕСКИХ ИЗМЕНЕНИЙ на прогнозируемую урожайность кукурузы в Соединенных Штатах. Кривые показывают, что максимальный урожай кукурузы может быть получен, если в период от цветения до образования семян температура будет несколько ниже, а дождевые осадки на 20% выше нормы. Линии соединяют точки, соответствующие таким сочетаниям значений температуры и дождевых осадков, при которых относительная урожайность одинакова. Например, кривая, отмеченная числом 80, соответствует климатическим условиям, при которых урожайность составит 80% среднего уровня. Подобные диаграммы составлены и по другим культурам для условий Соединенных Штатов и других стран. (Оценка произведена Университетом национальной обороны.)

Было бы неправильным считать, что эта проблема объясняется недостатком важных научных открытий и технических достижений. В перспективе их можно немало ожидать в таких областях, как гибридизация растений (например, пшеницы), повышение эффективности фотосинтеза, получение близнецов от животных, выведение засухоустойчивых растений и т.д. Но многие наблюдатели высказывают опасение, что разработка этих новшеств будет сдерживаться недостатком ассигнований на поисковые и экспериментальные работы и на программы, предусматривающие увеличение сельскохозяйственного производства на основе внедрения этих работ. Размер федеральных ассигнований на проведение исследований для сельского хозяйства начиная с 1975 г. стал значительно меньше по сравнению с периодом 1939—1965 гг., когда ежегодный прирост ассигнований на эти цели составлял 3,8%. В частном секторе расходы на исследовательские работы для сельского хозяйства определить трудно, но можно считать, что они не ниже государственных. Если даже расходы частного сектора увеличатся, они скорее всего будут направлены на разработку новшеств, которые можно запатентовать, т.е. на создание новых видов техники, химических веществ и пестицидов.

Прогнозируемые темпы внедрения достижений науки и техники имеют исключительное значение; именно научно-техническому прогрессу мы обязаны тем, что имели возможность удовлетворять спрос на продукты питания по сравнительно низким ценам и при этом не испытывать острой нехватки ресурсов. Если в ближайшее время не произойдет увеличения производительности труда в сельском хозяйстве или оно будет достигнуто не скоро, а ожидаемые потребности станут реальными, то мы вынуждены будем либо поднять цены и снизить экспорт, либо за счет других сфер производства увеличить ресурсы в сфере производства продовольствия.

Научно-техническая революция, переживаемая сельским хозяйством в течение последних трех десятилетий, сопровождалась огромными изменениями в его структуре. В период с 1950 по 1980 г. численность населения, проживающего на фермах, снизилась с 23 млн. человек до менее 8 млн., а средний размер фермерских хозяйств увеличился вдвое и в настоящее время стал равным более 160 га. Число мелких ферм еще довольно велико, но основная часть продукции производится постоянно укрупняющимися средними фермерскими хозяйствами. В прогнозах министерства сельского хозяйства США указывается, что если современные тенденции в изменении структуры сохранятся, то к 2000 г. на долю 50 тыс. крупнейших фермерских хозяйств будет приходиться 63% общего объема поставок сельскохозяйственной продукции по сравнению с 31% в 1974 г. Ожидается также, что специализация малых ферм станет более узкой, а разнообразие производимой ими продукции сократится по сравнению с прошлым периодом, когда мно-

гие небольшие американские фермы производили различные виды продукции и сами себя обеспечивали всем необходимым.

Проблема структуры имеет важное значение прежде всего потому, что с ней связаны социально-политические изменения в сельском хозяйстве, а в будущем она может оказать влияние и на производительность труда в этой сфере. Рост объемов производства сельскохозяйственной продукции часто сопровождается увеличением масштабов механизации и ирригации. Оба этих фактора усиливают зависимость сельского хозяйства от энергетических и других ресурсов, которые производятся вне ферм. Специализация небольших фермерских хозяйств ведет к локализации отдельных видов продукции; возникает региональная специализация, которая повышает зависимость производства от климатических условий, а сбыта продукции — от обеспечения энергией и работы транспорта.

Научно-технические достижения, которые позволили значительно повысить производительность труда в американском сельском хозяйстве, одновременно способствовали снижению устойчивости некоторых растений к болезням и вредителям. Современные методы селекции растений снижают генетическое разнообразие новых сортов. Так, более чем на 70% площади под кукурузой выращиваются сорта всего лишь шести инбредных линий. Узкая генетическая основа важнейших культур, например кукурузы, снижает их сопротивляемость болезням, вредным насекомым и климатическим воздействиям.

Слабая жизнестойкость выведенных сортов кукурузы привела к массовому заболеванию этого растения в США в 1970 г. мозаичной болезнью, вызванной одной из мутаций грибка *Helminthosporium maydis*. В результате было потеряно 15% урожая кукурузы. Хотя семеноводческие компании быстро внедрили резистентные семена и положение было вскоре исправлено, эпидемия, которая продолжалась в течение одного сезона, была воспринята как явление весьма опасное. Фермеры поняли, что стабильности урожая можно добиться только в условиях экологического равновесия, и что они принесли ее в жертву ради повышения производительности труда и снижения себестоимости производства сельскохозяйственной продукции за счет видовой унификации растений.

Существуют различные способы уменьшения нежелательных последствий при монокультуре. Улучшение коллекций зародышевых плазм позволяет быстро принимать соответствующие меры в случае возникновения эпидемий или заражения вредителями. Разработка действенных программ селекции растений и улучшение системы распознавания болезней и заражения вредителями также способствуют повышению жизнестойкости растений и их сохранению. На проведение научных исследований для сельского хозяйства следует выделять больше средств, а сами исследования должны быть ориентированы на решение актуальных проблем.

ИЗ МНОЖЕСТВА факторов, отрицательно влияющих на состояние американского сельского хозяйства, наибольшую трудность для анализа представляют, пожалуй, неблагоприятные климатические изменения. Эта трудность объясняется не столько ограниченным количеством имеющихся данных, сколько недостаточным пониманием причин колебаний климата. Сведения, которыми мы располагаем, свидетельствуют о том, что сельское хозяйство развивалось в необычно теплый период. Причем длительность таких теплых периодов в прошлом, подобных тому, в котором живем мы, продолжается не более 10 тыс. лет. Продолжительность нашего периода близка к концу этого срока. Климатологи расходятся во мнении, переживает ли наша планета длительный период постепенного потепления или похолодания. Как бы там ни было, для американского сельского хозяйства важнее годовые колебания погоды, а не длительные климатические процессы.

Анализ рассматриваемого явления еще более усложняется в связи с влиянием на климат человеческой деятельности. Вырубка лесов, образование пустынных территорий, производство и выброс в атмосферу различных отходов, содержащих кислоту, пылевые частицы, озон и двуокись углерода могут влиять на климат, а в конечном счете и на урожаи. Но пока влияние этих факторов на урожай не изучено. На помидоры и землянику, например, кислотосодержащие дожди оказывают благотворное влияние, а другим культурам, таким, как салат и шпинат, такой дождь может нанести вред.

В докладе Университета национальной обороны, опубликованном в 1980 г., была сделана попытка предсказать изменения урожая в зависимости от температурных условий и среднего уровня выпадения осадков. Специалистов сельского хозяйства, выступавших в качестве экспертов, просили оценить изменение урожайности при определенных сочетаниях условий в течение периода от цветения до образования семян различных культур в различных странах. Произведенные оценки урожая кукурузы в США показали, что максимальный урожай этой культуры может быть при условии, если температура в течение этого периода будет несколько ниже, а осадки на 20% выше нормы (*см. иллюстрацию на с. 57*). Урожай сои может быть максимальным, если температура не изменяется, а дождевые осадки на 40% выше нормы. Поскольку предсказания климатических изменений и их последствий остаются пока в сфере догадок, то лучше исходить из допущения, что иногда погода будет неблагоприятной.

СМОЖЕТ ЛИ американское сельское хозяйство еще в течение длительного времени производить в изобилии продукцию — вопрос, на который ответить не так просто. Лишь немногие эксперты предсказывают на будущее спад в сельскохозяйственном производстве и нехватку продовольствия для удовлетворения потребностей населения страны, которое сейчас ежегодно увеличивается менее чем на 1%. Неопределенность будущего положения объясняется отсутствием точных данных об уровне экспортных поставок в перспективе. Сейчас для удовлетворения экспортной потребности США вынуждены продавать другим странам урожай зерновых с каждого третьего гектара. До настоящего времени экспорт сельскохозяйственной продукции из США рассматривался как безусловное благо для страны. За счет экспорта повышаются доходы фермеров, обеспечивается приток иностранной валюты и сбыт постоянных излишков урожая зерновых, который временами, в том числе и сейчас, вызывает у американских фермеров серьезную озабоченность. Единственной причиной, существенно сдерживающей экспорт сельскохозяйственной продукции, до недавнего времени считалось его влияние на положение фермеров в странах-потребителях. Ввоз дешевого американского зерна подрывал конкурентоспособность местных производителей.

Но как только способность американских фермеров производить излишки стала меньше, то сразу стали выражаться сомнения в том, что неограниченное расширение экспорта отвечает национальным интересам. Следует ли способствовать зависимости других государств от американского зерна, продавая его по низким ценам, которые в конечном итоге сами США вряд ли смогут поддерживать? Нам говорят, что если американское сельское хозяйство достигнет предела своих возможностей, то это произойдет не только из-за стремления удовлетворить внутренние потребности в продовольствии. Расширение экспорта сельскохозяйственной продукции в будущем — это один из фундаментальных вопросов политики, который заслуживает такого же серьезного внимания, с каким в недавнее время рассматрива-

лось неблагоприятное положение в области обеспеченности сельского хозяйства необходимыми ресурсами.

Что касается будущего обеспечения сельского хозяйства ресурсами — землей, водой и энергией, то наибольшие трудности здесь, по-видимому, возникнут не столько из-за их физической нехватки, сколько из-за роста цен на них. В результате может усилиться процесс межрегиональных сдвигов в структуре производства растительных культур и продукции животноводства, равно как и в способах сельскохозяйственного производства. Повышение затрат на ирригацию и рост спроса на зерно могут, например, способствовать в животноводстве возврату к содержанию скота на пастбищах, как это было в прошлом веке. Спрос на землю тогда повысится и цены на скот и на землю возрастут.

Другие проблемы, как, например, эрозия, чрезмерная засоленность земель, климатические изменения и отставание с внедрением технологических новшеств,решать еще труднее. Если в перспективе повышение производительности труда в сельском хозяйстве будет происходить не столь быстро, как в последние годы, а умеренными темпами, которые были характерны для периода 1900—1950 гг., и ожидаемый экспортный спрос на американскую сельскохозяйственную продукцию при этом действительно окажется высоким и будет удовлетворяться, то мы можем стать свидетелями возврата к тем формам ведения сельского хозяйства, при которых увеличение производства зависит в основном не от повышения удельной урожайности земли, а от расширения ее площади. Снижение плодородия почв и неблагоприятные климатические изменения приобретут тогда бо́льшую важность, чем в случае, когда речь идет о повышении продуктивности.

Высокая степень монокультуры, характерная для американского сельского хозяйства, может представлять серьезную угрозу только в том случае, если не будут приняты простые превентивные меры, такие, как сохранение широкого разнообразия семян и создание здоровой и развитой индустрии по их производству. К сожалению, правительство не проявляет готовности взять под контроль решение этого вопроса.

В ЦЕЛОМ следует отметить, что предсказания, основанные на анализе тенденций в прошлом и настоящем, не свидетельствуют о наличии ситуации, близкой к кризису в американском сельском хозяйстве. В этих прогнозах представлены лишь наиболее вероятные ситуации, а не те, что менее вероятны, но все же возможны. Видимо, американская политика в области сельского хозяйства должна быть направлена на повышение роли страхования от наиболее неблагоприятных ситуаций — например, от погодных условий, снижающих продуктивность. Сумма страхования должна зависеть от рентабельности связанных с принятой стратегией превентивных мер. Возможными мерами могут быть: более рациональное территориальное распределение сельскохозяйственных земель, повышение разнообразия производимой в пределах района продукции, увеличение ассигнований на исследования для сельского хозяйства, создание частных и государственных запасов зерна, разработка программ, направленных на снижение воздействия угрожающих факторов, изменение правительственных программ, предусматривающих создание такой структуры сельского хозяйства, которая делает его неустойчивым в критических ситуациях, а также пересмотр цен на сельскохозяйственную продукцию, с тем чтобы они лучше отражали частные и общественные затраты на ее производство.

Исследования, связанные с разработкой этих мер, должны включать расчет расходов на борьбу с эрозией, уплотнением и засоленностью земель, определение влияния погодных условий на продуктивность сельского хозяйства, оценку расходов на введение в оборот потенциальных пахотных земель и оценку зависимости продуктивности от изменения территориальной структуры фермерских хозяйств. В большинстве случаев гораздо легче определить стоимость разработки и внедрения мероприятий профилактического характера, чем оценить возможную выгоду от мер по защите от чрезвычайных ситуаций.

При решении вопроса о том, какой должна быть политика в области сельского хозяйства, следует исходить не из оценки ожидаемых результатов исследований, а из стремления помочь выявить те области, которые до сих пор не изучены. В условиях современного положения экономики сельского хозяйства, которая задыхается от излишков, а фермеры обременены большими расходами и испытывают недостаток кредита, рассуждения по поводу того, как Соединенные Штаты должны противостоять возможной нехватке сельскохозяйственной продукции в будущем, выглядят по меньшей мере странными. Однако острота проблем, уже стоящих на повестке дня, не должна заслонять важность вопроса об улучшении организации в этом секторе экономики, чтобы он мог быть конкурентоспособным и процветать, пока существует наше общество.

Активная солнечная корона

Окружающее солнечный диск серебристо-белое сияние, которое раньше можно было наблюдать только при полных затмениях Солнца, сейчас является объектом постоянного исследования. Динамическая активность короны отражает сложные процессы взаимодействия вещества с магнитными полями

РИЧАРД ВОЛЬФСОН

Солнечная корона — внешняя разреженная оболочка Солнца, видимая лишь в моменты его полных затмений, представляет собой редкое и необычное зрелище. Почти каждый год в каком-нибудь из районов Земли можно наблюдать солнечное затмение, однако немногим удается за всю жизнь увидеть его более одного или двух раз. Из-за высокой яркости внешнего слоя Солнца, называемого фотосферой, корона обычно не заметна, и только благодаря тому, что видимые размеры Луны и Солнца почти одинаковы, ее удается иногда наблюдать. Будь Луна чуть меньше, полных солнечных затмений не было бы вообще, а окажись она чуть больше, внутренняя область солнечной короны была бы скрыта от наших глаз.

За прошедшее десятилетие наши знания о короне, как и о само́м Солнце, значительно пополнились. Это стало возможным благодаря новым приборам, установленным не только на Земле, но и на ракетах, спутниках и космических кораблях. Как показывают наблюдения, корона постоянно находится в состоянии истечения, время от времени прерываемого бурными всплесками. «Ветер» из частиц, выбрасываемых короной, пронизывает всю Солнечную систему; изменения скорости и интенсивности этого ветра свидетельствуют об активности короны. Такое представление о короне и ее проявлениях позволяет рассматривать Солнце уже не как компактный объект в центре Солнечной системы, а как целостное образование гигантских размеров, простирающееся за пределы орбит наиболее удаленных планет. Корона — это расположенная недалеко от нас огромная лаборатория для исследования астрофизических процесссов. Ее изучение будет способствовать расширению наших представлений о различных явлениях, например о процессах, протекающих в термоядерных реакторах. Вполне возможно, что мы узнаем много нового и о взрывах сверхновых звезд в глубинах Вселенной.

На обычных фотографиях корона выглядит как серебристо-белое гало вокруг заслоненного Луной солнечного диска. Светимость гало обусловлена в основном рассеянием солнечного света корональными электронами. Из каждого миллиона фотонов, покидающих солнечную фотосферу, только один рассеивается в короне. Поэтому яркость фотосферы в миллион раз превышает яркость короны. Прозрачность короны указывает на то, что составляющий ее газ чрезвычайно разрежен. По земным стандартам корона — это отличный вакуум: максимальная концентрация ее частиц находится в пределах 10^8 — 10^9см $^{-3}$.

Спектроскопические исследования короны показывают, что ее состав примерно такой же, как и у Солнца: в основном это водород и гелий со следами более тяжелых элементов. Эти исследования позволяют также судить о температуре короны. Известно, что нагревание газа приводит к ионизации его атомов, т.е. потере ими одного или более электронов. Газ в короне ионизован гораздо сильнее, чем в фотосфере, температура которой близка к 6000 К. Атомы водорода и гелия в короне, по существу, лишены электронов, а атомы более тяжелых элементов теряют бо́льшую часть своих электронов. Высокая степень ионизации свидетельствует о том, что корона чрезвычайно горяча. Кроме того, ее спектральные линии уширены, что говорит о хаотичности движения корональных частиц с высокими скоростями. Результаты спектральных наблюдений показывают, что температура короны лежит в пределах 1 — 2 млн. К.

Высокая температура короны долгое время представляла серьезную теоретическую проблему. Согласно второму началу термодинамики, более холодное тело не может самопроизвольно отдавать тепло более горячему. Поэтому поток тепла от Солнца, как бы он ни был велик, не может нагреть корону до температуры выше 6000 К. Это положение носит столь фундаментальный характер, что десятки лет астрономы, изучающие Солнце, с недоверием относились к полученным спектроскопическим данным. Выдвигались различные хитроумные гипотезы, например, что корональные ионы возникают в недрах Солнца, где температура достаточно высока, или рождаются непосредственно в короне в результате ядерных реакций. В 40-х годах концепция горячей короны стала общепринятой. Она позволила теоретикам быстро продвинуться вперед в описании общей физической природы короны.

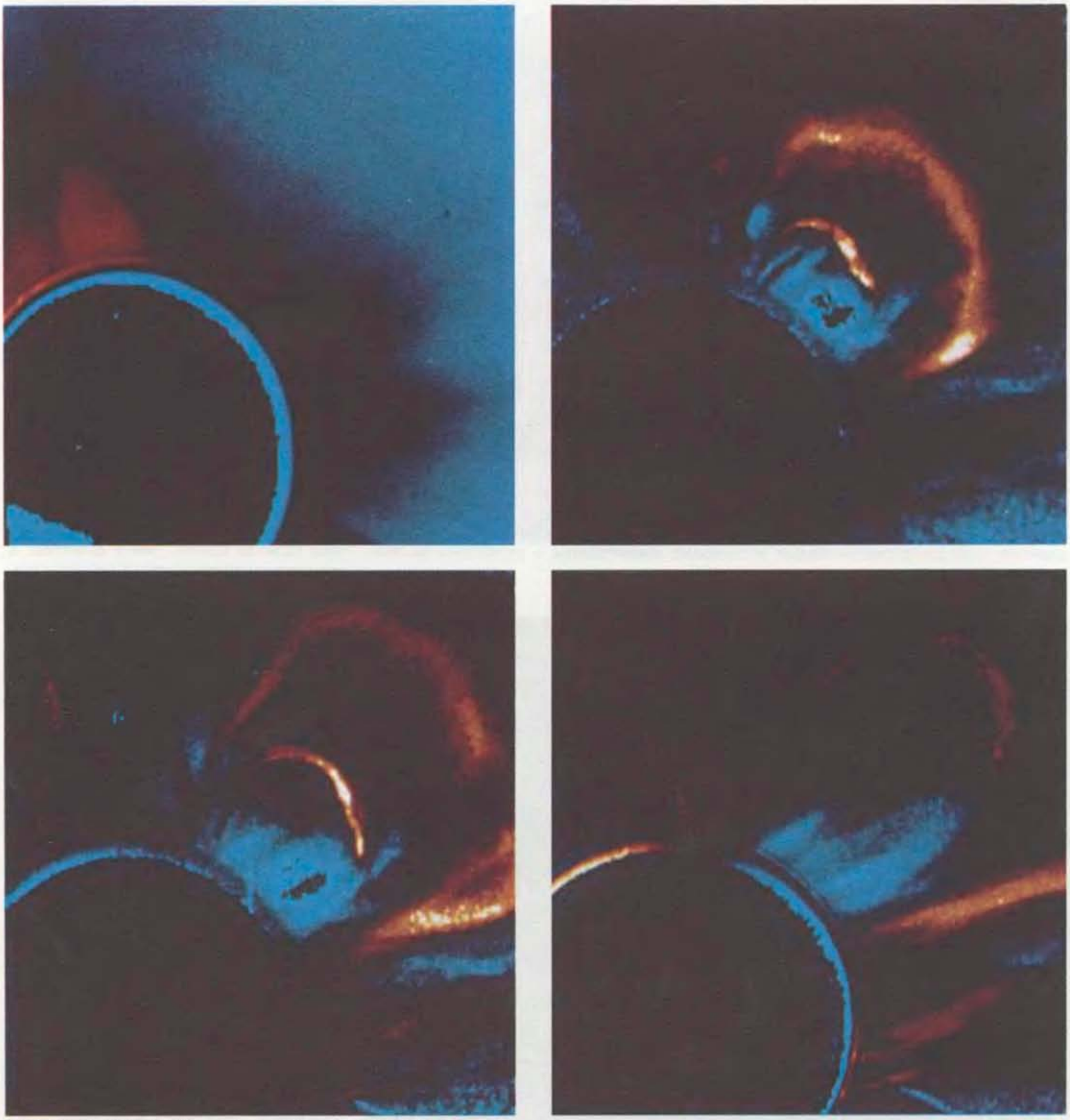
Если второе начало термодинамики исключает возможность переноса тепловой энергии от фотосферы к короне, то как происходит ее нагрев? Наиболее вероятно, что он обусловлен прямым переносом механической энергии. Согласно современным представлениям, механическая энергия переносится в корону из конвективного слоя Солнца, находящегося под фотосферой, в форме мощных акустических или магнитогидродинамических волн.

Акустические волны — это механические возмущения газа, которые проявляются в виде изменений давления. Магнитогидродинамические волны, которые появляются в ионизованном газе, — это возмущения при взаимодействии ионизованного вещества и магнитных полей. Хотя общая концепция механического нагрева короны волновыми процессами утвердилась достаточно прочно, для построения детальных моделей, согласующихся с наблюдениями, потребуются усилия многих исследователей. В результате, по-видимому, удастся связать физические свойства короны и нижележащих слоев, что позволит ближе подойти к пониманию физической природы Солнца в целом.

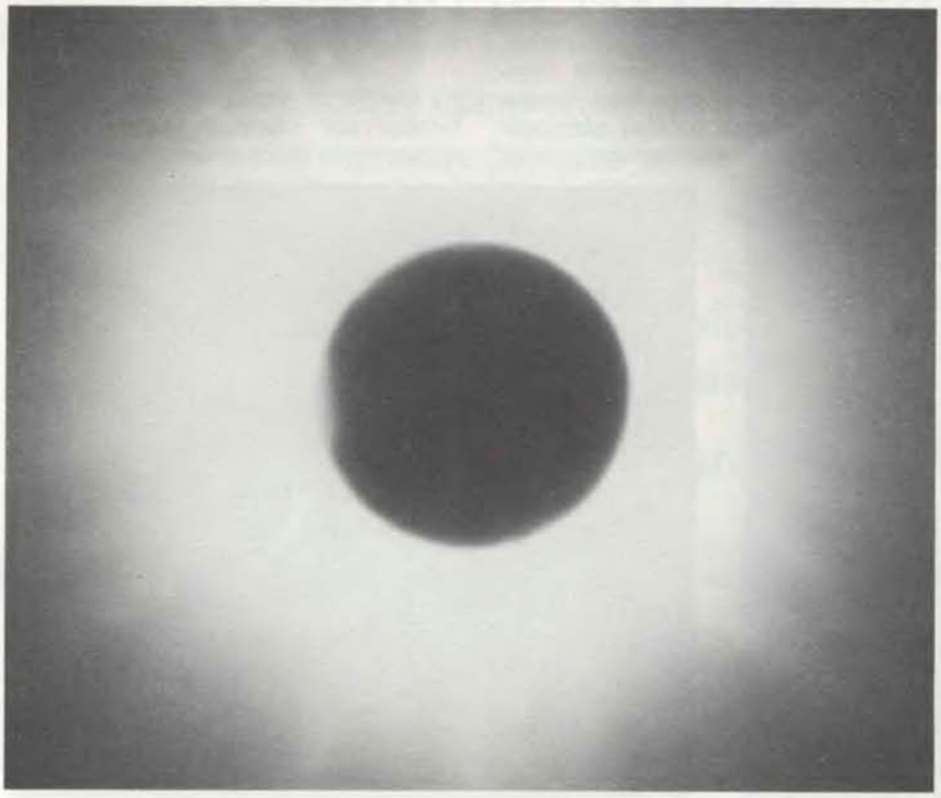
Итак, корона — это область с высокой температурой и низкой плотностью, окружающая менее нагретую, более плотную и в большей степени изученную фотосферу. (Фотосферу отде-

ляет от короны тонкий промежуточный слой, называемый хромосферой.) Большинство обычных фотографий показывают, что корона далеко не симметрична, вблизи полюсов Солнца она менее яркая и менее протяженная. Заметные выпуклости вдали от полюсов говорят о сложности ее структуры. Эти отклонения от симметрии меняются от затмения к затмению. Наиболее симметричные короны наблюдаются в периоды максимумов солнечных пятен в ходе 11-летнего цикла активности Солнца.

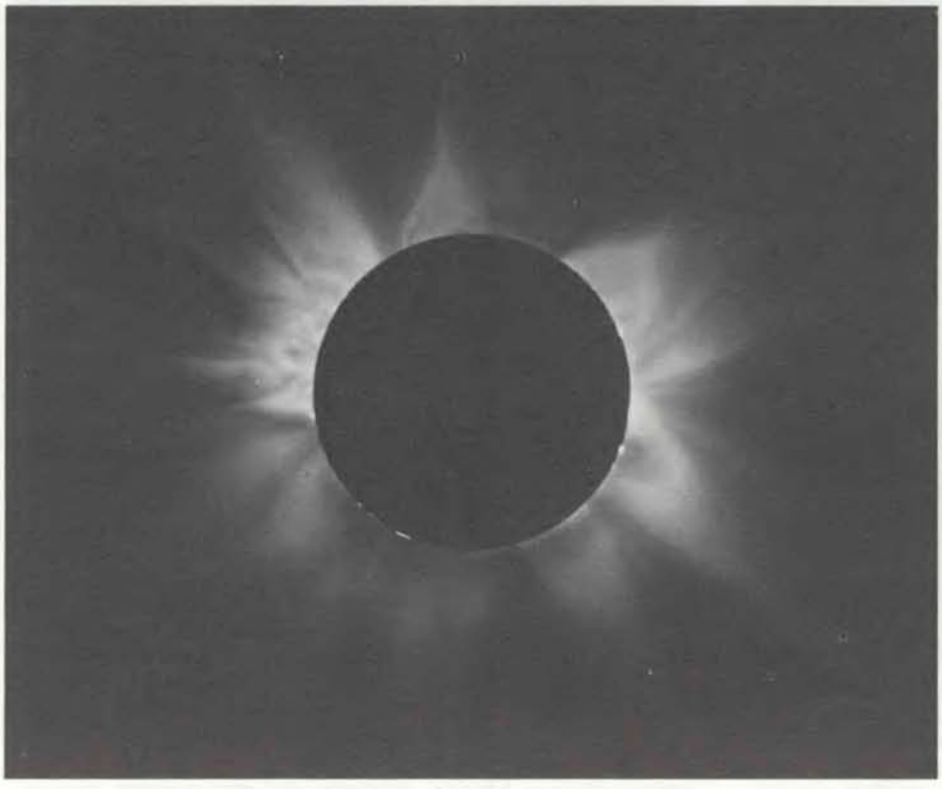
Судить о структуре короны по обычным фотографиям весьма трудно из-за ее яркой внутренней обла-



БЫСТРО ИЗМЕНЯЮЩАЯСЯ ПЕТЛЯ В СОЛНЕЧНОЙ КОРОНЕ, обнаружена в апреле 1980 г. при наблюдении с помощью коронографа, установленного на спутнике SMM. Коронограф — это телескоп, в котором специальный круглый экран заслоняет яркий свет солнечной фотосферы (видимой поверхности Солнца), позволяя наблюдать в миллион раз менее яркую корону. Коронограф для спутника SMM разработан на Высокогорной обсерватории Национального центра исследований атмосферы (Боулдер, шт. Колорадо). Свет от короны фокусируется на фотокатоде телевизинной камеры, полученное изображение передается на Землю в цифровой форме. Корона до начала быстропротекающего процесса изображена на первом из четырех последовательных снимков (цвета произвольные). Три других снимка получены методом цифрового вычитания: яркость в каждой точке соответствует изменению интенсивности света относительно уровня, зарегистрированного для первого изображения. *Вверху справа*: корональная петля имеет размер, примерно равный диаметру Солнца. *Внизу слева*: спустя 25 мин петля, расширяясь со скоростью примерно 300 км/с, достигает края изображения; вслед за ней начинает подниматься вверх протуберанец. *Внизу справа*: спустя 20 мин видно только одно основание петли. Неизвестно, связаны между собой корональная петля и протуберанец или оба явления вызваны неустойчивостью магнитного поля Солнца.



ОБЫЧНАЯ ФОТОГРАФИЯ ЗАТМЕНИЯ СОЛНЦА, сделанная в Индии 16 февраля 1980 г. группой специалистов во главе с Дж. Пасаховом (Вильямсовский колледж). Неразличимые здесь детали показаны на нижнем снимке. Тем не менее даже на этой фотографии видно, что корона симметрична — особенность, связанная с 11-летним циклом солнечных пятен.



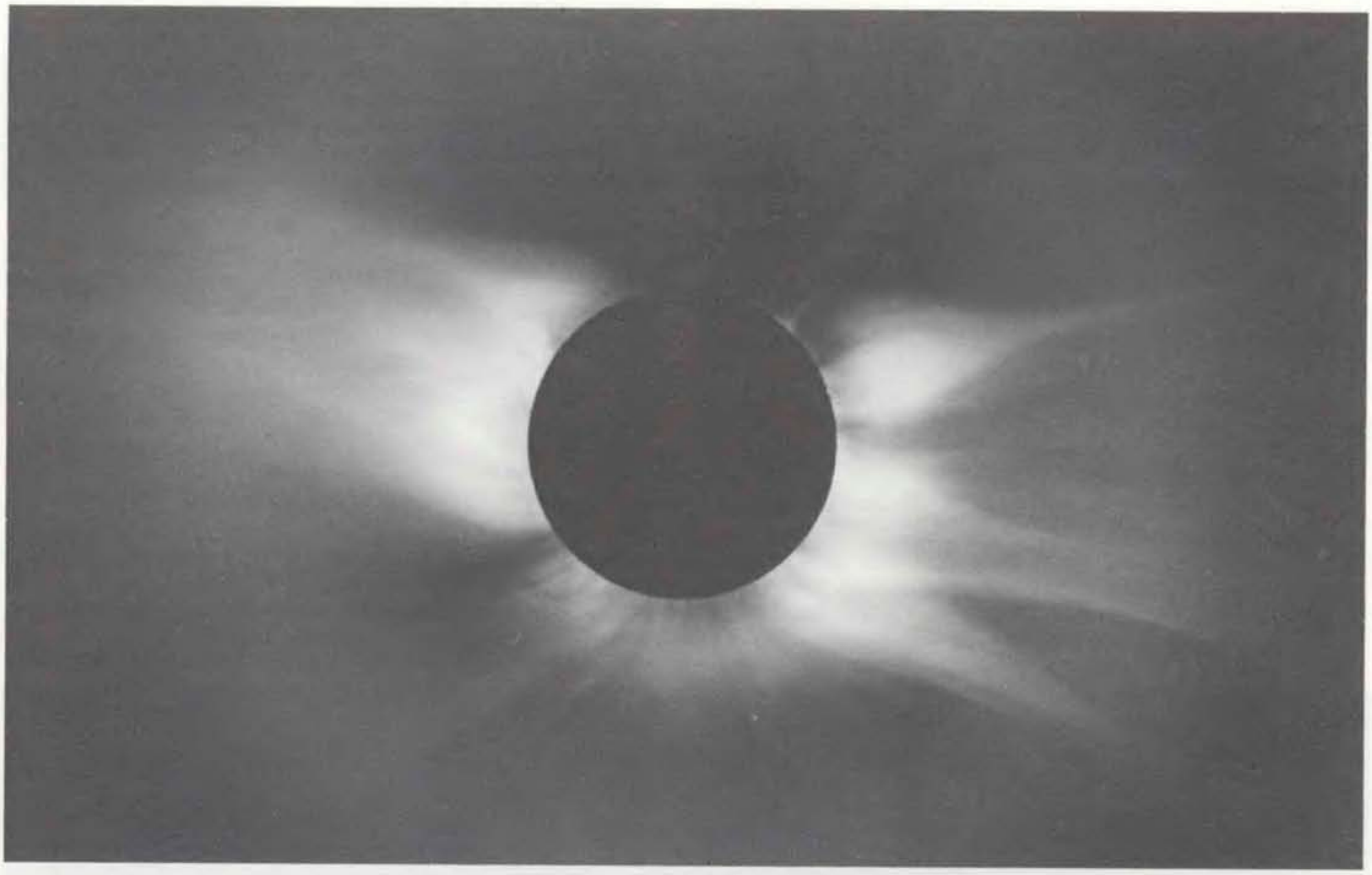
ТОНКИЕ ДЕТАЛИ СТРУКТУРЫ КОРОНЫ во время затмения 16 февраля 1980 г. удалось сфотографировать с помощью фильтра с оптической плотностью, меняющейся вдоль радиуса. Такой фильтр позволяет выявить детали корональной структуры от уровня фотосферы до расстояния в несколько солнечных радиусов. Яркие пятна выше диска Луны — протуберанцы. Хорошо видны корональные лучи (стримеры). Снимок сделан в Индии группой исследователей из Высокогорной обсерватории.

сти, приводящей к передержке при фотографировании. Самым успешным способом решения этой проблемы является фотографирование короны через специальный фильтр, плотность которого меняется вдоль радиуса таким образом, что свет в центре изображения ослабляется в большей степени, чем по краям, и тем самым компенсируется падение яркости короны. С помощью таких фильтров удается проследить корональную структуру на расстоянии многих солнечных радиусов от фотосферы.

На фотографии, полученной с использованием такого фильтра, солнечная корона предстает во всем своем причудливом великолепии. Длинные тонкие лучи (стримеры) вытягиваются на несколько миллионов километров. Ближе к солнечному диску видны яркие петли и арки из коронального вещества. На некоторых фотографиях заметны также прямые узкие корональные лучи. В периоды максимумов солнечной активности эти детали структуры короны окружают весь солнечный диск, в периоды же минимумов активности они наблюдаются в пределах более низких солнечных широт. В последнем случае в короне видны большие «дыры» на обоих полюсах Солнца, где она гораздо слабее, чем у экватора. В полярных «дырах» часто заметны слабые полярные «щеточки». От каждого полюса расходится своеобразный пучок этих щеточек, ослабевая и становясь невидимым на расстоянии примерно одного солнечного радиуса. Основные детали структуры — лучи, петли и дыры — настолько преобладают, что ряд исследователей полагают, что корона состоит только из них и не имеет подструктуру, которую можно было бы назвать «спокойной короной».

Вид короны существенно меняется от одного затмения к другому. Даже обычные фотографии затмений различаются настолько, что можно без труда определить, в момент какого затмения производилась съемка. Снимки короны через радиальный фильтр сильно отличаются друг от друга. Помимо общих изменений, связанных с циклом солнечной активности, для каждого затмения характерна специфическая конфигурация корональных лучей, петель и других деталей структуры. Серия фотографий, сделанных в различные моменты в течение одного затмения, свидетельствует о динамичности короны, структурные детали которой изменяются за время, гораздо меньшее, чем интервалы между затмениями.

Для изучения динамики короны необходимо вести наблюдения даже в отсутствие затмения, что практически невозможно в оптическом диапазоне, в котором плотная фотосфера оказывается в миллион раз ярче разреженной



СИММЕТРИЧНАЯ КОРОНА. Сфотографирована через радиальный фильтр во время затмения 30 июня 1973 г.; время затмения пришлось почти на минимум 11-летнего цикла солнечных пятен. В такие периоды корональные лучи сконцентрированы в экваториальной области, а вблизи полюсов Солнца появляются «дыры» — области пониженной активности. Снимок сделан в Кении исследователями из Высокогорной обсерватории.

короны. Однако корона гораздо горячее фотосферы, поэтому испускаемое ею электромагнитное излучение высокой энергии практически отсутствует в спектре последней. Имея температуру порядка 2 млн. К, газ короны оказывается мощным источником рентгеновского излучения. Если не учитывать рентгеновское излучение, связанное с солнечными вспышками и другими активными областями, то более холодная фотосфера оказывается пренебрежимо слабым источником такого излучения. Поэтому корону можно изучать с помощью рентгеновского телескопа.

Такие наблюдения приходится выполнять в космосе, поскольку земная атмосфера непрозрачна для рентгеновских лучей. Первые рентгеновские телескопы, устанавливавшиеся на космических аппаратах в начале 1970-х годов, позволили получить грубые снимки короны, на которых различимы яркие области с высокой плотностью или температурой. Сопоставление снимков затмений, сделанных с Земли, с рентгеновскими снимками, сделанными в космосе примерно в то же время, позволило установить соответствие между областями, яркими в рентгеновском диапазоне, и оптически наблюдаемыми деталями структуры короны.

Метод получения рентгеновских изображений быстро развивался в последнее десятилетие. В середине 1970-х годов астронавты на борту орбитальной обсерватории «Скайлэб» получили множество рентгеновских снимков короны. Киномонтаж таких снимков наглядо показывает, как изменяются и перемещаются корональные структуры при вращении Солнца. Полученные на «Скайлэбе» рентгеновские снимки позволили установить, что корональные дыры существуют не только на полюсах, но и по всей короне, хотя имеют меньшие размеры. С помощью усовершенствованной аппаратуры недавно были выполнены наблюдения короны в рентгеновском диапазоне с более высоким разрешением.

Дополнительную информацию о короне может дать анализ ультрафиолетового излучения, источником которого являются ее высокоионизованные тяжелые элементы. В этой области спектра излучение фотосферы также пренебрежимо мало. Данные об ультрафиолетовом излучении особенно ценны тем, что определенные спектральные линии соответствуют только узким интервалам температур; это позволяет как бы зондировать различные слои короны, хромосферы и переходной области между ними.

Исследование короны в радиодиапазоне возможно с помощью наземных радиотелескопов, способных регистрировать радиосигналы, связанные с характерными колебаниями ионизованного коронального газа. По частоте таких колебаний удается судить непосредственно о плотности короны, различные области которой зондируются путем изменения частоты, на которую настроен радиотелескоп. Методом радиоинтерферометрии с помощью нескольких радиотелескопов можно проследить за эволюцией отдельных корональных особенностей. В Калгуре (Австралия) установлен инструмент, принадлежащий Государственной организации по научным и промышленным исследованиям, названный радиогелиографом. Он состоит из 96 связанных между собой радиотелескопов, предназначенных для исследования Солнца.

Создание искусственного затмения — наиболее прямой подход к исследованию короны между полными затмениями Солнца. В принципе даже

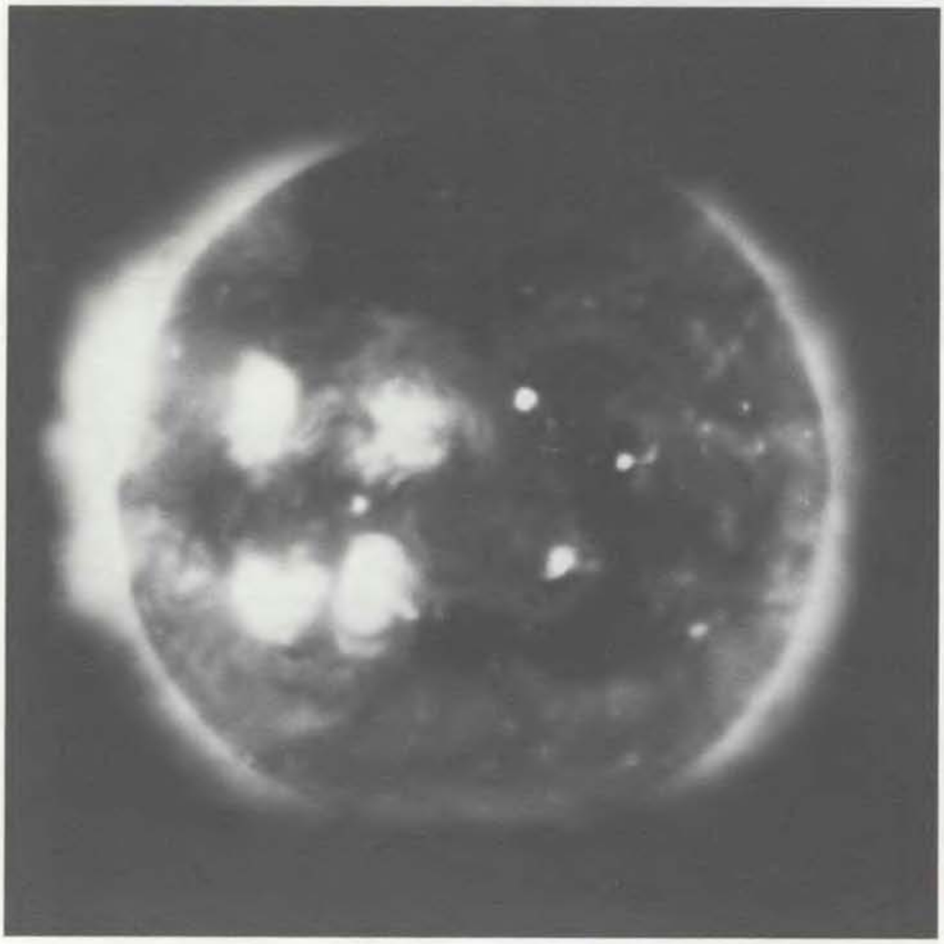
большой палец на вытянутой руке мог бы заслонить свет солнечной фотосферы. Однако этот простой прием оказывается непригодным из-за того, что рассеиваемый земной атмосферой солнечный свет буквально «забивает» свет короны. Тем не менее за пределами атмосферы и даже на высоких горах оказывается возможным наблюдать солнечную корону, искусственно экранируя свет от фотосферы. Разработанный для этой цели инструмент получил название коронографа. Это небольшой телескоп-рефлектор, снабженный круглым экраном, заслоняющим солнечный диск, подобно Луне, во время затмения. Однако приходится добавлять еще несколько различных экранов и диафрагм для максимального ослабления рассеянного света и дифракционных эффектов, искажающих изображение короны. Выделенный свет короны либо сразу регистрируется, либо проходит через поляризаторы или фильтры для детального анализа.

Коронографы, появившиеся около 50 лет назад, устанавливаются на горных вершинах, на стратостатах и на борту космических аппаратов. Коронограф на «Скайлэбе» позволил открыть быстроменяющиеся структуры в короне, которые, по-видимому, выбрасывают корональное вещество в межпланетную среду. Коронографы имеются на борту нескольких беспилотных космических аппаратов. Основные наземные коронографы размещены на Сакраменто-Пик (Нью-Мексико, США), на Пик-дю-Миди (Франция) и на горе Мауна-Лоа (Гавайские о-ва). В 1979 г. орбитальный коронограф Морской исследовательской лаборатории (США) зарегистрировал событие, никогда не наблюдавшееся ранее, — столкновение кометы с Солнцем. В новейших коронографах для регистрации изображения используется не фотопленка, а цифровые электронные детекторы, сигналы с которых обрабатываются на ЭВМ, что позволяет получать контрастные изображения с высоким разрешением.

Одними из лучших являются снимки короны, полученные с помощью коронографа на спутнике SMM (Solar Maximum Mission), запущенном в феврале 1980 г. примерно в период максимума солнечной активности. На спутнике была установлена аппаратура для исследования Солнца в диапазоне длин волн от видимого света до гамма-лучей. За шесть с небольшим месяцев спутник SMM передал множество данных, но затем в результате серии неисправностей в системе ориентации на Солнце, его вращение вышло из-под контроля, поэтому программа исследований с помощью этого спутника, рассчитанная на два года, была прервана. Однако спутник SMM — первый космический аппарат, сконструированный с учетом возможности его ремонта в космосе. Он снабжен специальным крюком, а приборы размещены в сменных блоках. Программа полета космического корабля «Шаттл», запланированного на апрель 1984 г., предусматривает сближение со спутником SMM, во время которого астронавт, выйдя в космос, остановит его неуправляемое вращение. Механическая рука «Шаттла» перенесет спутник в грузовой отсек корабля, и после устранения в нем неисправностей он будет возвращен на прежнюю орбиту. Восстановленный спутник SMM должен возобновить наблюдения в период минимума солнечной активности; полученные данные можно будет сравнить с данными, переданными спутником в период максимума солнечной активности.

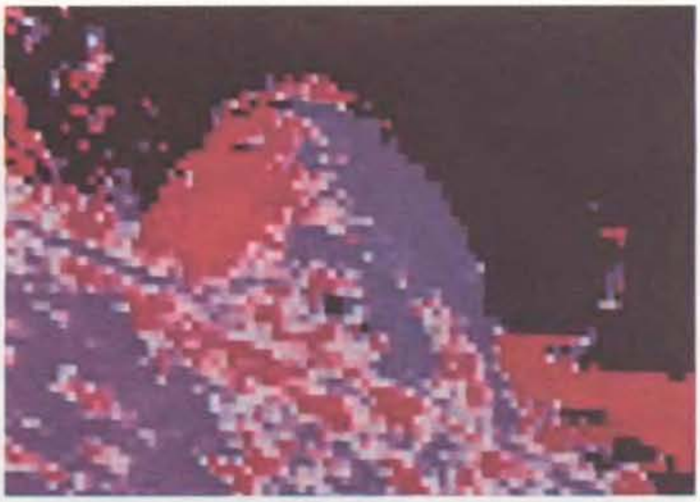
Коронограф на спутнике SMM сконструирован группой исследователей под руководством Л. Хауза из Высокогорной обсерватории Национального центра исследований атмосферы (США). В этом приборе свет от короны проходит через взаимозаменяемые фильтры и поляризаторы и фокусируются на фотокатоде телевизионной камеры. С камеры данные передаются на Землю в цифровой форме и записываются на магнитную ленту. Каждое передаваемое изображение содержит свыше миллиона бит информации и соответствует одному квадранту солнечной короны. Принятые наземными станциями изображения обрабатываются ЭВМ и преобразуются в цветное телевизионное изображение. С помощью ЭВМ наблюдатель может выделить интересующие его области короны и получить увеличенные изображения отдельных деталей.

До временного перерыва в работе осенью 1981 г. коронограф спутника SMM передал около 30 тыс. изображе-

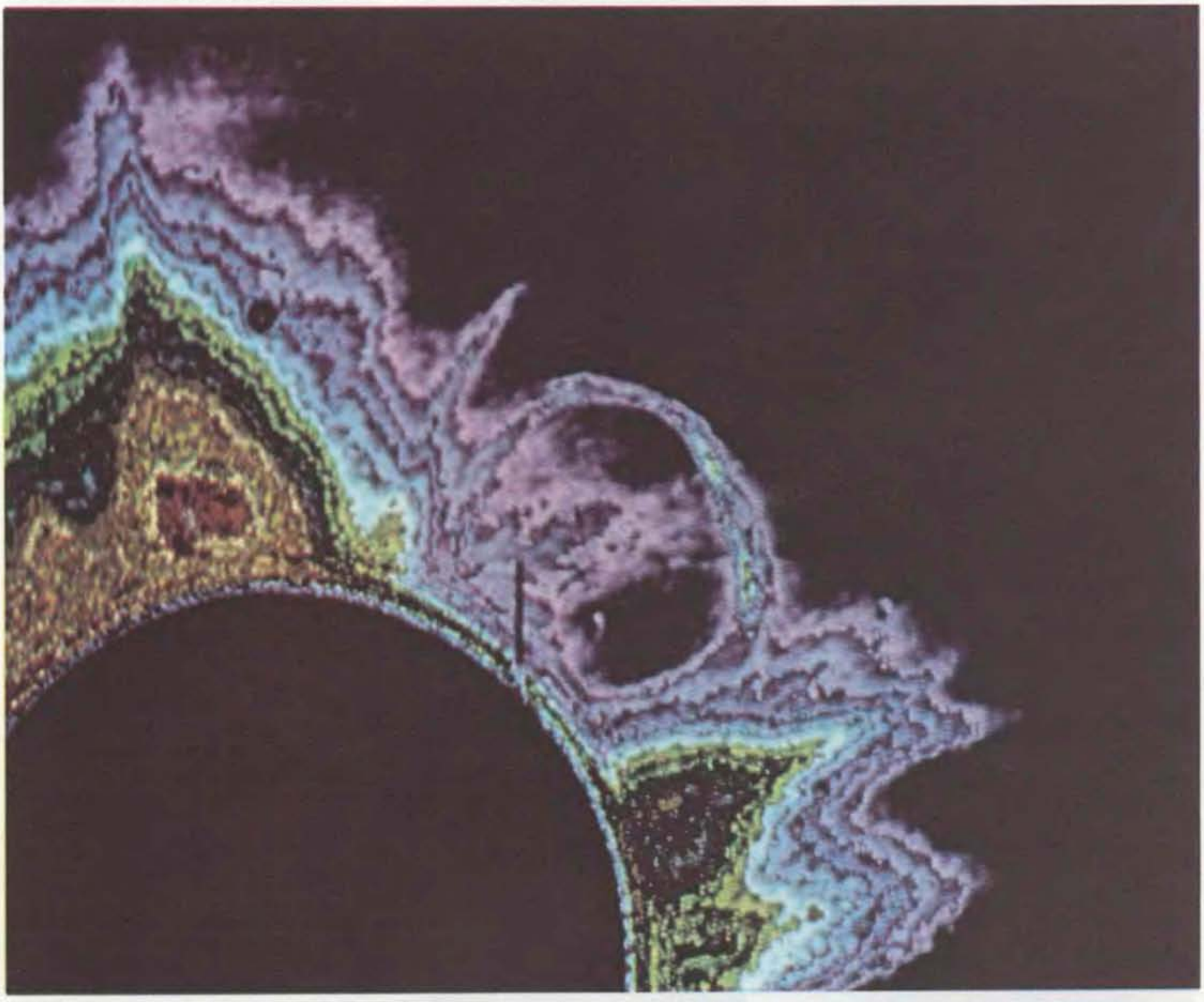


РЕНТГЕНОВСКИЙ СНИМОК СОЛНЦА. Корона, температура которой достигает примерно 2 млн. К, является источником рентгеновского излучения. Фотосфера, имеющая температуру 6000 К, слишком холодна, чтобы быть источником этого излучения. Снимок сделан астронавтами с орбитальной станции «Скайлэб» 18 сентября 1973 г., примерно через два с половиной месяца после получения снимка, показанного на с. 63. Яркие «пятна» — это корональные области повышенной температуры или плотности, соответствующие корональным лучам или петлям. Темные области — это корональные дыры. Видна также дыра в экваториальной области. Подобные снимки Солнца показывают, что корональные дыры образуются не только в полярных районах и что они связаны с высокоскоростными потоками частиц, выбрасываемых из Солнца. Рентгеновский телескоп, с помощью которого получен снимок, создан фирмой American Science and Engineering, Inc.





ПЕТЛЯ У ГРАНИЦ ФОТОСФЕРЫ, поддерживаемая сильным локальным магнитным полем, имеет температуру около 100 000 К. Изображения получены в ультрафиолетовых лучах с длиной волны 1548,19 Å, испускаемых трехкратно ионизованными атомами углерода. Снимок слева сделан в узком спектральном интервале с центральной длиной волны, указанной выше. Изображение справа — «скоростеграмма» той же петли, полученная разделением излучения на две составляющие — меньшую и бо́льшую относительно центральной длины волны. Составляющая с меньшей длиной волны (синяя) испускается веществом, приближающимся к наблюдателю (ультрафиолетовому спектрометру на спутнике SMM), составляющая с большей длиной волны (красная) — веществом, удаляющимся от него. Такие скоростеграммы хорошо иллюстрируют динамичный характер внешних слоев Солнца.

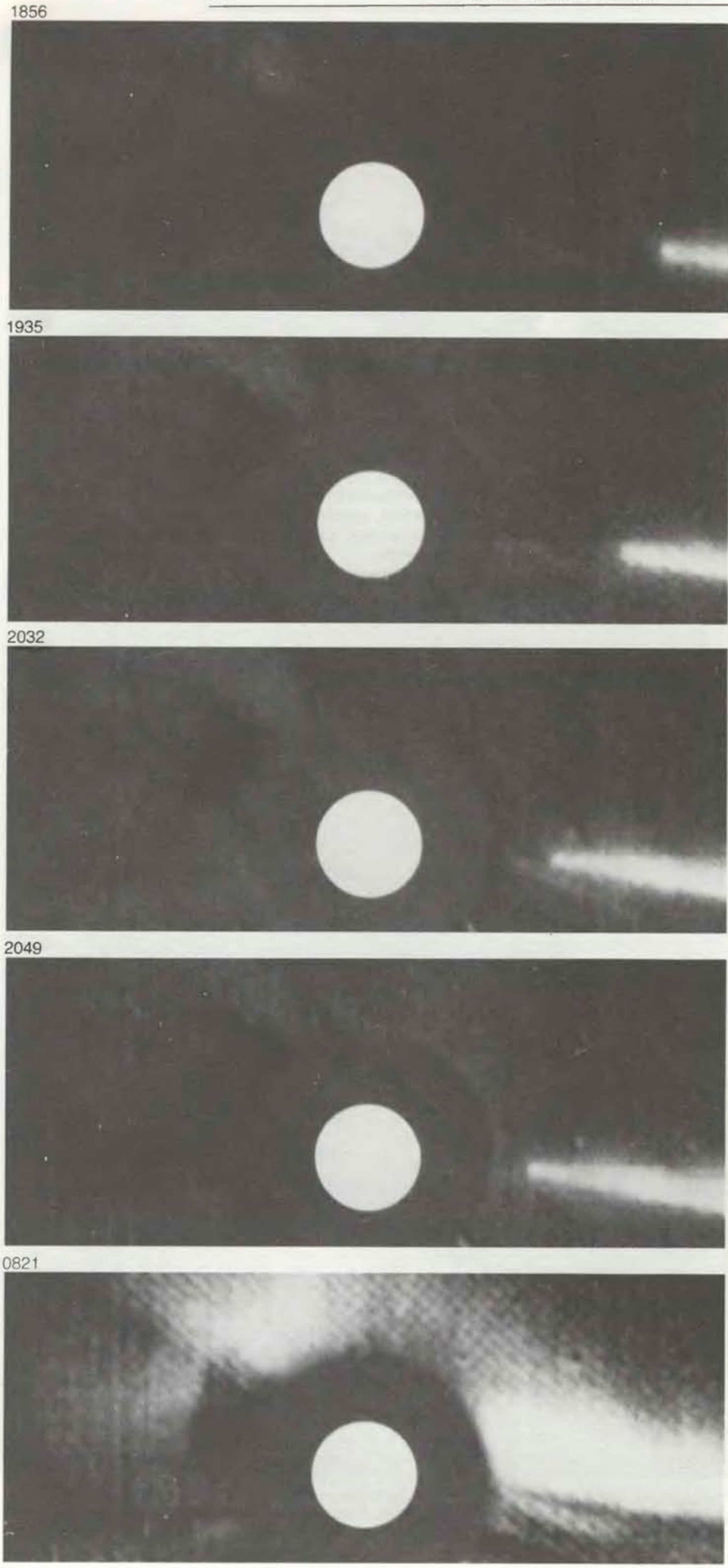


КОРОНАЛЬНЫЕ ЛУЧИ И ПОДНИМАЮЩИЙСЯ ПРОТУБЕРАНЕЦ сняты 14 апреля 1980 г. при помощи коронографа Высокогорной обсерватории, установленного на спутнике SMM. Контрастность увеличена за счет приема «эффективного обозначения контура» — выделение деталей изображения при помощи ЭВМ. Различные цвета соответствуют разной плотности короны. На снимке — высококонтрастное изображение явления, 4 фазы которого приведены на с. 61.

ний короны. Р. Иллинг из Высокогорной обсерватории сделал киномонтаж серии изображений, полученных в апреле 1980 г., что позволило впервые увидеть корону как неспокойную, постоянно изменяющуюся динамическую структуру. Хотя основные ее детали остаются различимыми в течение нескольких дней или более, они явно изменяются за это время. Меньшие детали структуры появляются и изменяются еще быстрее. Случайные быстропротекающие процессы «развиваются» в короне в течение нескольких часов. Коронограф SMM позволил зарегистрировать один из таких процессов, представляющий собой внезапное расширение корональной петли и ее разрастание в изогнутую вовне арку со скоростью порядка нескольких сотен километров в секунду.

КАКОВА физическая природа деталей корональной структуры? Как может горячий разреженный газ образовывать такие удивительные узоры? Почему одни корональные детали существуют несколько дней, а другие изменяются быстро и неожиданно? Ни сила тяготения, характеризующаяся сферической симметрией, ни силы, связанные с вращением Солнца, не могут придать короне ее необычный облик. Разгадка происхождения корональных структур кроется в их формах. Отдельные корональные петли, арки и полярные щеточки образуют рисунки, напоминающие расположение железных опилок вблизи магнита. Подобно тому как опилки выстраиваются вдоль силовых линий магнитного поля, так и в корональном газе проявляется структура солнечного магнитного поля. Именно магнитные силы ответственны за детали корональных структур, их разнообразие и изменчивость.

Подобно магнитному полю Земли и многих других небесных тел, магнитное поле Солнца, по-видимому, обусловлено конвективными движениями и вращением проводящего электричество вещества внутри него. По причинам, до сих пор не выясненным, маг-

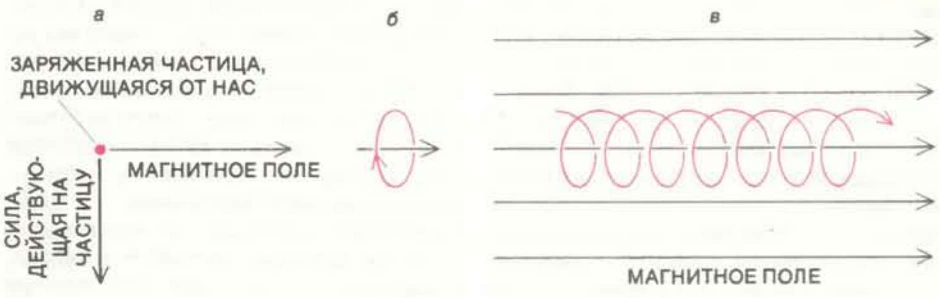


КОМЕТА СТАЛКИВАЕТСЯ С СОЛНЦЕМ. Впервые такое событие удалось запечатлеть при помощи орбитального коронографа Научно-исследовательской лаборатории Военно-морского флота США. Снимок вверху сделан в 18 ч 56 мин мирового времени 30 августа 1979 г. Само Солнце полностью закрыто искусственной «луной» коронографа; белый диск соответствует размеру и положению Солнца. Комета, принадлежащая семейству комет, близко подходящих к Солнцу, до этого не была зарегистрирована на Земле. Яркие области, видные на нижнем снимке, существовали в течение примерно 24 ч после столкновения. С 1979 г. наблюдались еще два подобных столкновения.

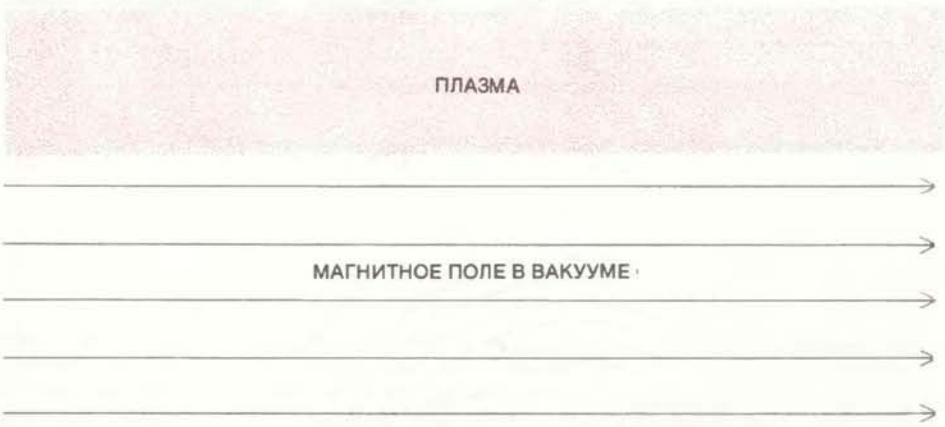
нитное поле Солнца очень изменчиво. Примерно каждые 11 лет его полярность меняется. Наиболее выраженный характер полюсы имеют в периоды минимума солнечных пятен, в это время общая конфигурация поля проста и напоминает поле магнитного диполя, каковыми являются, например, стержневой магнит и Земля. Дипольная структура поля хорошо проявляется в полярных щеточках, заметных в короне в периоды минимума солнечных пятен. В периоды максимума пятен магнитное поле хаотично; вместо четко выраженных северного и южного полюсов наблюдаются сильные локальные магнитные поля, часто связанные с солнечными пятнами. Корональные петли и арки, особенно многочисленные в это время, обрисовывают границы локальных полей.

Взаимодействие коронального газа и магнитного поля подчиняется основным законам электромагнетизма. Изучение магнитных явлений началось с изучения магнитных свойств черных металлов, таких, как железо. Известно, что магнитное поле создается движущимися электрическими зарядами. В магнитных материалах такими зарядами являются электроны, обращающиеся вокруг атомных ядер; именно этим движением электронов и обусловлены магнитные свойства этих материалов. Частицы ионизованного газа, такого, как корона, заряжены и, следовательно, взаимодействуют с магнитным полем. Эти электромагнитные взаимодействия делают ионизованный газ настолько отличным от обычных газов, что он получил специальное название — плазма. Плазму нередко называют четвертым состоянием вещества. На такой относительно холодной планете, как наша Земля, плазма встречается редко. Однако для многих объектов Вселенной это состояние вещества, по-видимому, является основным.

На заряженную частицу в магнитном поле действует сила, зависящая от ее заряда, скорости а также напряженности и направления поля. Направление этой силы перпендикулярно направлению движения частицы и направлению магнитного поля. Она максимальна, когда частица движется под прямым углом к направлению поля, и равна нулю при ее движении параллельно направлению поля. Поскольку магнитная сила всегда перпендикулярна направлению движения частицы, то она может изменить только направление, в котором движется частица, но не ее скорость. При движении частицы параллельно направлению поля эта сила равна нулю, поэтому магнитное поле оказывает на нее воздействие только тогда, когда ее скорость имеет составляющую, перпендикулярную полю. В этом случае заряженные частицы движутся



ЗАРЯЖЕННЫЕ ЧАСТИЦЫ В КОРОНЕ ведут себя так же, как и заряженные частицы в магнитном поле. Действующая на них сила перпендикулярна магнитному полю и направлению их движения. На положительно заряженную частицу, движущуюся перпендикулярно плоскости рисунка (*а*) «от нас», действует сила, направленная вниз, если магнитное поле направлено слева направо. Под действием этой силы частица начинает двигаться по окружности (*б*). Если скорость частицы имеет составляющую вдоль поля, то она будет двигаться по спирали (*в*). Отрицательно заряженная частица будет двигаться в направлении, противоположном указанному на рисунке.



ПЕРЕМЕЩЕНИЮ ПЛАЗМЫ под прямым углом к магнитному полю препятствует оказываемое этим полем давление. Вакуум, «пронизанный» силовыми линиями магнитного поля, граничит с находящейся под давлением плазмой. Проникновению плазмы в эту область препятствует давление магнитного поля. Но эту конфигурацию нельзя считать устойчивой. Небольшое возмущение в плазме или магнитном поле может вызвать ее резкое изменение.



МАГНИТНАЯ ПЛАВУЧЕСТЬ проявляется в области более сильного магнитного поля, находящегося между областями с более слабым полем. Полное давление в каждой области равно сумме давлений магнитного поля и плазмы. Если все области имеют одинаковую температуру, то меньшее давление соответствует и меньшей плотности. На область с меньшей плотностью сила тяготения оказывает меньшее действие и она стремится «всплыть». Этим можно объяснить выбрасывание вещества петель в межпланетную среду.

по спиралям, вокруг силовых линий. Говорят, что они «вморожены», или «привязаны» к магнитному полю, подобно тому как связаны с ниткой нанизанные на нее бусины. Частицы могут перемещаться только вдоль направления поля. Неоднородность поля и столкновения между частицами делают их менее «привязанными» к полю, однако для разреженной плазмы, подобной солнечной короне, допущение, что ее частицы «привязаны» к магнитному полю, является вполне приемлемым.

ЗАРЯЖЕННЫЕ частицы, составляющие корону, образуют своеобразный газ, который, как и газ земной атмосферы, оказывает давление. Вообразим, что корональный газ стремится под действием собственного давления проникнуть в область, где имеется магнитное поле. Поскольку частицы плазмы не могут свободно двигаться поперек силовых линий поля, то поле активно противодействует давлению газа. Количественный анализ показывает, что магнитное поле действительно оказывает давление на ионизованный газ, причем это давление пропорционально квадрату напряженности поля. Поведение коронального газа гораздо сложнее поведения земной атмосферы, поскольку его динамика определяется как газовым давлением, так и магнитным.

Корональная плазма отнюдь не пассивно реагирует на солнечное магнитное поле. Поскольку корона разрежена и высокоионизована, она оказывается отличным проводником электричества и может поддерживать весьма сильные электрические токи. Последние в свою очередь генерируют магнитные поля, изменяющие первоначальное поле. Этим взаимодействием короны с магнитным полем Солнца обусловлено богатое разнообразие ее свойств, которое не характерно для непроводящего газа.

Широкий диапазон корональных явлений значительно усложняет попытки теоретического объяснения поведения короны. Для полного понимания корональных процессов необходимо получение непротиворечивого решения сложной системы уравнений, описывающей процессы формирования магнитным полем корональной структуры с учетом корональных токов, изменяющих само поле. Кроме того, нельзя пренебрегать действиями тяготения, вращения и газового давления. Не удивительно, что теория, охватывающая весь комплекс корональных явлений, еще не разработана.

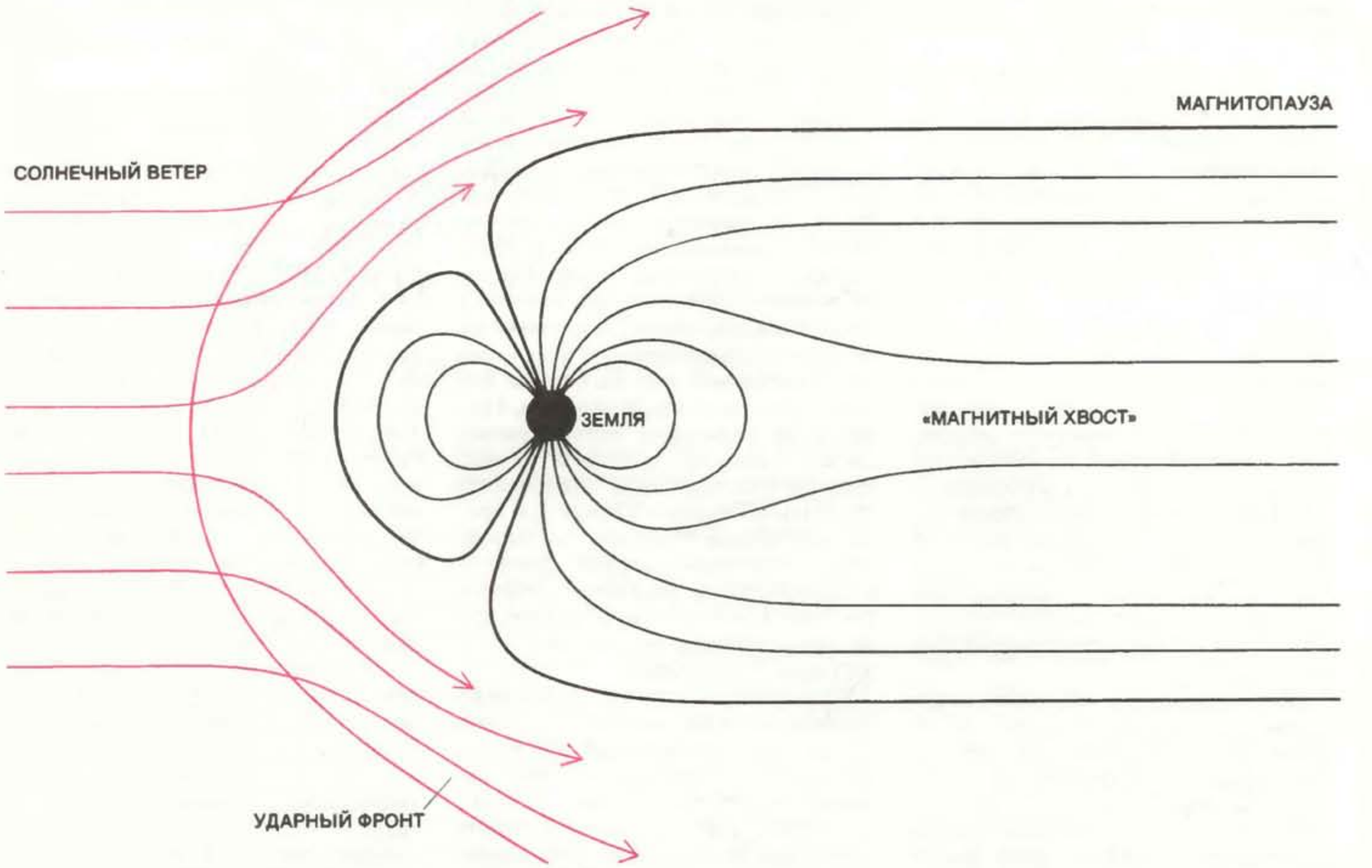
Первым важным шагом в этом направлении явилась разработка математической модели Дж. Ньюманом и Р. Коппом из Высокогорной обсерватории. Модель хорошо описывает развитие корональных лучей в экваториальной области Солнца. Ньюман, Тян Е из Национального управления по исследованию океанов и атмосферы (США) и другие исследователи усовершенствовали модели экваториальных лучей, что привело к более полному пониманию их структуры. В настоящее время предпринимаются попытки установить, как связаны корональные лучи с ближайшими к ним корональными дырами. Разрабатываемая мною модель описывает образование лучей в высоких солнечных широтах.

Открытия, сделанные за последнее десятилетие, стимулировали теоретические исследования динамики солнечной короны, в особенности быстропротекающих корональных процессов. Работа, выполненная недавно Б. Лоу из Высокогорной обсерватории и автором настоящей статьи, позволяет предположить, что быстрый подъем корональных петель обусловлен магнитной плавучестью. Подобно тому как на погруженную в воду щепку действует выталкивающая сила, так и на область более сильного магнитного поля в короне действует выталкивающая сила со стороны окружающего газа. Теоретические модели показывают, что корональная петля может внезапно потерять устойчивость и подняться в короне либо вследствие хаотического движения фотосферного газа у основания магнитной петли, либо в ответ на бурное возмущение в фотосфере, например солнечную вспышку.

Поскольку корональная плазма и магнитное поле связаны друг с другом, плазма не может двигаться перпендикулярно направлению поля, не увлекая за собой силовые линии поля. Поэтому возмущения в плазме могут привести к вздутию силовых линий. Это вздутие может перемещаться вдоль линий,



БЫСТРОПРОТЕКАЮЩИЙ ПРОЦЕСС РАСШИРЕНИЯ петли описан с помощью математической модели, недавно разработанной автором этой статьи. Кривые представляют последовательные положения петли в магнитном поле. Вершина петли поднимается по мере увеличения расстояния между ее основаниями в результате, например, солнечной вспышки или хаотического движения газа у поверхности Солнца. Согласно модели, при расширении петли скорость ее подъема все более превышает скорость расхождения оснований. Такая картина свидетельствует о неустойчивости: относительно небольшое возмущение на поверхности Солнца может быть причиной быстрых изменений в структуре короны. В модели учтены эффекты газового давления, магнитного поля и тяготения. Для упрощения математического описания предполагалось, что магнитное взаимодействие происходило в тонких слоях, ограниченных изображенными кривыми.



ДАВЛЕНИЕ СОЛНЕЧНОГО ВЕТРА удерживает магнитное поле Земли в ограниченной области, называемой магнитосферой. Эта область сжата в направлении к Солнцу и вытянута в противоположном направлении в длинный «магнитный хвост». Магнитопауза — граница между солнечным ветром и магнитосферой — является своеобразным барьером, препятствующим попаданию на Землю основной массы частиц солнечного ветра. Однако вблизи полюсов Земли частицы могут перемещаться вдоль силовых линий магнитного поля, попадая в верхнюю атмосферу и вызывая полярные сияния. Перед магнитосферой образуется ударный фронт, как при полете сверхзвукового самолета.

подобно тому как бежит изгиб вдоль натянутой струны. Такое перемещающееся возмущение представляет собой магнитогидродинамическую волну, переносящую энергию в форме отклонений как плазмы, так и магнитного поля от состояния равновесия, характеризующегося наименьшей энергией. Как отмечалось выше, перенос энергии из лежащего под фотосферой конвективного слоя через фотосферу и хромосферу непосредственно в корону может быть обусловлен магнитогидродинамическими волнами. В виде ударных волн они рассеивают свою энергию в короне, разогревая ее до температуры в 2 млн. К.

Почему изучение динамики короны так важно? Оказывает ли корональная активность какое-либо прямое влияние на Землю? Изучение короны позволит глубже исследовать сложное поведение горячего ионизованного газа в магнитном поле. Полученные результаты пригодятся при исследовании гораздо более далеких и загадочных объектов, таких, как пульсары, квазары и активные галактики. Если мы лучше будем понимать процессы, происходящие в короне, это поможет нам использовать термоядерные реакции в качестве источника энергии, так как наиболее перспективные направления в этой области связаны с проблемой удержания горячей плазмы магнитными полями.

ЕЩЕ БОЛЕЕ значительным представляется тот факт, что Земля «погружена» в солнечную корону. В конце 1950-х годов стало общепризнанным, что именно высокой температурой короны обусловлен солнечный ветер — поток солнечного газа, скорость которого вблизи Земли достигает 400 км/с. Солнечный ветер — это истечение короны в межпланетную среду.

Поскольку корональная плазма «привязана» к солнечному магнитному полю, то области короны, окруженные магнитными петлями, не могут порождать солнечный ветер. В самом деле, магнитные петли обоими концами связаны с поверхностью Солнца, поэтому корональный газ из этих петель не истекает в межпланетную среду. Газ попадает в нее вдоль силовых линий магнитного поля Солнца, простирающихся в бесконечность. В настоящее время считают, что солнечный ветер, в особенности его высокоскоростные потоки, образуется в основном в корональных дырах. Быстропротекающие корональные процессы могут «проложить» еще один путь для вещества, покидающего Солнце. Во многих таких явлениях первоначально замкнутая петля становится неустойчивой, поднимается и увеличивается в размерах до тех пор, пока ее силовые линии не размыкаются, приводя к выбросу вещества в межпланетную среду. Недавние наблюдения выявили существование своеобразных «корональных снарядов» — результата процесса конденсации холодного плотного вещества, которые быстро ускоряются при прохождении через корону, увеличивая истечение вещества из Солнца.

Магнитное поле Земли защищает ее от прямого потока солнечного ветра, который, однако, искажает это поле. В направлении к Солнцу поле Земли сжато, а в противоположном направлении вытянуто в виде длинного хвоста. Область магнитного поля Земли, охваченная солнечным ветром, называется магнитосферой. Как и солечная корона, магнитосферы Земли и других планет — это своего рода лаборатории для изучения взаимодействия плазмы с магнитным полем.

Магнитосфера представляет собой препятствие на пути солнечного ветра. Поскольку его скорость превышает скорость звука или магнитогидродинамических волн, ни один сигнал, также

являющийся возмущением, не может быть передан от магнитосферы к набегающему потоку так, чтобы солнечный ветер мог плавно обогнуть это препятствие. Ситуация аналогична движению самолета быстрее звука или лодки быстрее волн на поверхности воды. В этих случаях не происходит плавное обтекание препятствия, как это имело бы место при движении со скоростью, меньшей скорости звука или скорости распространения другого вида волн. В результате возникает ударная волна, и такие свойства среды, как давление, плотность и скорость, резко меняются. Ударная волна, создаваемая сверхзвуковым самолетом, порождает «звуковой удар»; от быстродвижущейся лодки разбегается носовая волна. Перед земной магнитосферой образуется обширный ударный фронт на расстоянии около 60 000 км от Земли. Космический аппарат, пересекающий этот фронт, может собрать богатую информацию о структуре ударной волны в плазме.

При неизменном солнечном ветре магнитосфера и ударный фронт оставались бы относительно друг друга в фиксированном положении, определяемом равновесием между полным давлением набегающего солнечного ветра и давлением магнитного поля Земли. Однако ветер — это продолжение короны, которая подвержена значительным изменениям как в пространстве, так и во времени. По мере вращения Солнца различные пространственные нерегулярности короны — такие, как корональные дыры и магнитные петли, — преобразуются во временные вариации солнечного ветра, достигающие Земли. Динамические явления на Солнце (вспышки или быстропротекающие процессы в короне) вызывают изменения солнечного ветра.

В результате изменения давления солнечного ветра ударный фронт и магнитосфера смещаются, сжимая магнитное поле Земли и изменяя его напряженность. Изменения магнитного поля возбуждают электрические токи в ионосфере — проводящем слое верхней атмосферы. Это в свою очередь может привести к внезапным нарушениям дальней радиосвязи. Такие же токи могут возбуждаться в телефонных линиях и линиях электропередач — как воздушных, так и подземных, — а также в земле. Поэтому сильные возмущения на Солнце могут искажать показания компасов, нарушать связь и даже подачу электроэнергии. В это время частицы высокой энергии, двигаясь по спиральным траекториям вдоль магнитных силовых линий, попадают в полярные области Земли и вызывают полярные сияния.

Высокоскоростные потоки солнечного ветра вырываются из корональных дыр. По мере того как такой поток «наталкивается» на движущийся впереди поток ветра, обладающий меньшей скоростью, переходная зона между этими потоками уменьшается, а сам переход становится более резким. К тому времени, когда солнечный ветер достигнет орбиты Юпитера, удаленного от Солнца в пять раз дальше, чем Земля, эти переходные зоны превратятся в ударные волны, которые сталкиваются с магнитным полем Юпитера. Магнитосфера Юпитера заполняет такой большой объем, что, несмотря на огромное расстояние от Земли, она по размерам казалась бы больше Солнца, если бы мы могли ее видеть. Под давлением солнечного ветра ее размеры могут изменяться почти в три раза всего за несколько дней. Космические аппараты «Пионер» и «Вояджер», прошедшие вблизи Юпитера, последовательно пересекли ударный фронт и магнитопаузу, в результате удалось установить, что эти две области быстро перемещаются взад и вперед, то вздуваясь, то сжимаясь.

Если солнечная корона столь явно проявляет себя вблизи Юпитера, то как далеко она простирается? На каком расстоянии Солнце еще способно оказывать воздействие? В какой-то точке постепенно уменьшающееся давление солнечного ветра должно стать слабее очень малого давления межзвездного газа и магнитного поля Галактики. Эта область перехода указывает положение гелиопаузы, т.е. границы физической протяженности Солнца. Теоретические расчеты, выполненные около десяти лет назад, позволили предположить, что эта граница проходит между орбитами Юпитера и Сатурна. Однако космический аппарат, ныне приближающийся к орбите Нептуна, не обнаружил никаких признаков гелиопаузы. По нынешним представлениям, солнечный ветер простирается за орбиту самой далекой планеты Солнечной системы.

За гелиопаузой должен существовать ударный фронт, порожденный движением всей Солнечной системы в межзвездной среде. Из-за непостоянства солнечной короны гелиосфера —область, лежащая внутри гелиопаузы, — может сжиматься и расширяться подобно воздушному шару. От термоядерной «печи» в центральной области Солнца до гелиопаузы, включая конвективную зону, фотосферу, хромосферу, корону и солнечный ветер, — все это составляющие Солнца. Человечество занимает привилегированное положение в пределах этого гигантского небесного тела — наша планета расположена достаточно далеко от его ядра, так что на ней создались благоприятные условия для развития жизни и существования человека, и достаточно близко к нему, чтобы можно было проникнуть в тайны короны.

Книги

Ида Томпсон. КРАТКИЙ ОПРЕДЕЛИТЕЛЬ ИСКОПАЕМЫХ СЕВЕРНОЙ АМЕРИКИ.

THE AUDUBON SOCIETY FIELD GUIDE TO NORTH AMERICAN FOSSILS, by Ida Thompson. Alfred A. Knopf, Inc. ($ 12.50)

НА СУПЕРОБЛОЖКЕ этой небольшой книжки, которую очень удобно носить с собой в кармане, изображен переливающийся перламутром аммонит. Эту раковину размером в десять сантиметров нашли много лет назад в Южной Дакоте; она хранится в музее Пибоди в Йельском университете, там ее и сфотографировали. Фотографии раковин — главное в определителе. Все, кто работал над книгой, от японских типографских рабочих, оформителя и фотографа до автора, палеонтолога из Университета Ратжерса, заслуживают самых высоких похвал.

«Определитель» — первоклассная книга; она, несомненно, займет свое место в библиотеке натуралиста-любителя. Справочный аппарат не уступает иллюстрациям. Чтобы идентифицировать находку, нужно отыскать ее контур в таблице силуэтных изображений, затем обратиться к указанным цветным таблицам и найти сходные формы, после чего круг поисков сужается, если определить, к какому геологическому периоду относится находка — для этого даны подробные карты, — и, наконец, остается выбрать наиболее подходящее изображение. Текст не только помогает искать решение, но и обогащает ваши познания об ископаемых растениях и животных. Автор, как видно, счел, что ледниковые ландшафты и граниты Новой Англии, а также докембрийский Квебек недостаточно богаты ископаемыми, и карты этих районов не приводятся. Быть может, у жителей этих мест книга вызовет желание путешествовать — пусть хотя бы по музеям. Образцы, фотографии которых вошли в книгу, хранятся в одной из четырех лучших коллекций музеев восточных штатов; все они определены и датированы. Книгу отличает тщательность и продуманность исполнения.

Ядовитые лягушки

В Колумбии индейцы-охотники отравляют иглы для духовых трубок сильными ядами — алкалоидами, которые добывают из маленьких лягушек. И сами алкалоиды, и эволюционная биология ядовитых лягушек представляют большой интерес

ЧАРЛЗ У. МАЙЕРС И ДЖОН У. ДЭЛИ

НА НАШЕЙ планете живет столько разнообразных ядовитых существ, что пройдет еще немало времени, прежде чем химики и фармакологи сумеют изучить состав всех ядов, вырабатываемых микроорганизмами, растениями и животными. Фундаментальные исследования в этой области предоставляют широкое поле деятельности для экологов и эволюционистов. Кроме того, эти работы могут привести к открытию новых лекарственных средств и других полезных человеку соединений. Особенно интересные объекты встречаются в богатых флорой и фауной тропических областях. Например, растения рода *Strychnos*, распространенные в тропиках Нового Света, — источник яда стрихнина и алкалоидов группы кураре. Индейцы-охотники отравляют кураре свои стрелы и иголки для духовых трубок, а анестезиологи применяют это вещество как мышечный релаксант. Другой пример — особые лягушки, которые выделяют один из сильнейших ядов, известных в животном мире. Яд лягушек также используется индейцами. В настоящее время его интенсивно изучают и медики, и биологи. Очень интересны и особенности биологии ядовитых лягушек.

Ядовитые лягушки и их неядовитые сородичи относятся к семейству Dendrobatidae. Его представители обитают в Южной Америке и на юге Центральной Америки, где занимают самые разнообразные экологические ниши: одни виды живут вблизи воды — по берегам речек и ручьев, другие — на суше в дождевых лесах гор и низменностей, некоторые лесные виды проводят бóльшую часть жизни на деревьях. Есть и такие виды, которые обитают на открытых сухих пространствах, довольствуясь влажностью затененных участков почвы под низкорослыми растениями. Несмотря на столь различные требования к условиям обитания, у всех лягушек-дендробатид есть общие экологические особенности, которые отличают их от остальных лягушек. Лягушки-дендробатиды активны только в дневное время суток. Они откладывают икру прямо на землю во влажных местах и заботятся об икринках до вылупления головастиков, после чего лягушка-"нянька" носит головастиков в буквальном смысле приклеенными к своей спине, пока не найдет подходящего водоема.

Семейство Dendrobatidae включает 130 видов, в настоящее время их объединяют в 4 рода. Это мелкие лягушки длиной от 1 до примерно 5 см. В то время как у большинства лягушек самцы, как правило, мельче самок, во многих популяциях дендробатид самцы и самки приблизительно одинаковой величины. Крупным самцам свойственно ярко выраженное территориальное поведение, они весьма агрессивны. Много времени и энергии уходит у них на обозначение границ своей территории, они привлекают партнерш продолжительными криками и, кроме того, активно конкурируют и дерутся с другими самцами. Самцы территориальных видов — необыкновенно шумные создания. Их стрекотанье, чириканье, трели (очень похожие на звуки, издаваемые насекомыми) составляют существенную часть дневного шума некоторых тропических лесов. Представьте себе двух крохотных лягушек, вполне умещающихся на одном листе дерева, которые затеяли состязание и ожесточенно стрекочут друг на друга. Такое зрелище может вызвать улыбку, но не стоит забывать, что это проявление полового поведения, которое обеспечивает успешное размножение, — серьезное дело с эволюционной точки зрения. У некоторых видов дендробатид самки также проявляют агрессивность и активно защищают свою территорию.

С ОТНОСИТЕЛЬНО крупными размерами самцов у дендробатид связаны также необычные формы их полового поведения. У большинства лягушек ухаживание заключается в том, что самец влезает на спину более крупной самки и сжимает свою подругу позади ее передних лап или же перед задними. Самцы дендрóбатид обхватывают самок не за туловище, а за голову, прижимая тыльные стороны пальцев к подбородку самки. По сравнению с обычной позой самец продвинут вперед; это положение самца наилучшее для оплодотворения откладываемых самкой икринок: если бы самец такого размера уцеплялся ближе к заднему концу тела самки, его семенная жидкость просто попадала бы на землю. У некоторых видов дендробатид во время ухаживания "объятия" не происходит, а икринки оплодотворяются в ходе довольно сложного маневрирования, в котором принимают участие как самец, так и самка.

Разные виды дендробатид откладывают икринки в лесную подстилку, в расселины под камнями, на затененные низко нависающие над землей листья или в тропические растения из семейств бромелиевых и ароидных. В то время как лягушки,размножающиеся в воде, откладывают сотни и тысячи икринок, у дендробатид кладка невелика: в ней может насчитываться всего одна-две икринки (у самых мелких представителей семейства), а вообще их редко больше 30 — 40. Во многих популяциях небольшой объем кладки частично компенсируется тем, что дендробатиды размножаются на протяжении всего года. Некоторые особи могут давать потомство каждый месяц. Статистический материал по этому вопросу дали наблюдения за парой лягушек вида *Dendrobates tricolor*, более двух лет живущих в Американском музее естественной истории. Эти лягушки спаривались 2 — 4 раза в месяц, кладка появлялась в среднем каждые 10 дней и насчитывала от 12 до 30 икринок. Даже при такой сравнительно высокой интенсивности размножения самка откладывала ежегодно только около 600 икринок.

Отложенные икринки один из родителей либо охраняет постоянно, либо периодически навещает. Во время таких визитов лягушка смачивает икринки водой, которую приносит, повидимому, в своем голосовом мешке. Время от времени лягушка-нянька перемешивает массу икринок задними

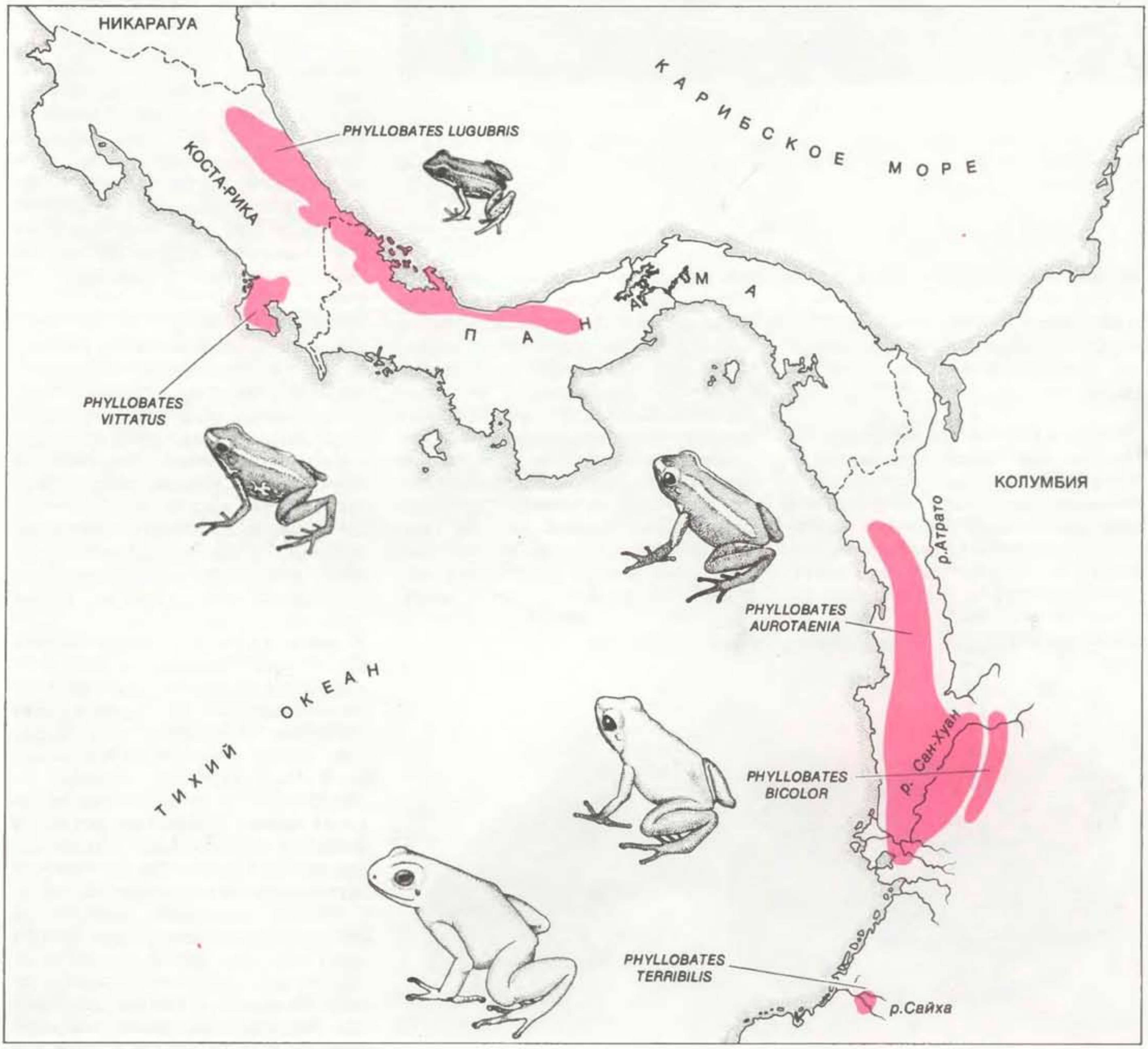


ЯДОВИТЫЕ ЛЯГУШКИ рода *Phyllobates* изображены на этом рисунке в 1,7 натуральной величины. Две лягушки *в верхнем ряду слева и в центре* — центральноамериканские виды *P. lugubris* и *P. vittatus*. Полосатая лягушка *вверху справа — P. aurotaenia. В среднем ряду* — два экземпляра *P. bicolor*, у представителей этого вида лапы окрашены в другой цвет, чем туловище. *В нижнем ряду* — лягушка *P. terribilis. P. aurotaenia*, *P. bicolor* и *P.terribilis* обитают в Колумбии: их ядом индейцы пропитывали свои охотничьи иглы. Лягушки *P. terribilis* — самые ядовитые из этих трех видов. Неполовозрелые особи *P. terribilis* (*слева*) полосатой расцветкой напоминают своих центральноамериканских сородичей (*вверху*), однако с возрастом полосы исчезают. Взрослые особи *P. terribilis* (*справа*) могут быть желтыми, оранжевыми или светло-зелеными, обычно окраска верхней и нижней поверхности их тела одинакова. Лягушки *P. terribilis* самые храбрые; по сравнению с другими они выделяют меньше яда, зато их яд примерно в 20 раз сильнее, чем у прочих, поэтому к ним опасно даже прикасаться. Рисунок Д. Денниса.

лапами — это стимулирует вылупление и помогает появившимся головастикам вползти на родительскую спину. Там они прикрепляются к пятнам особой слизи, которую выделяют находящиеся в коже лягушки железы. У головастиков поверхность брюшка плоская или слегка вогнутая, что помогает им держаться на спине у няньки. У некоторых видов головастики просто сидят на ней, у других слизь клейкая, и головастики буквально приклеиваются к спине взрослой особи. Головастики могут оставаться на лягушке-няньке от нескольких часов до недели и более (у разных видов этот период различен). За это время они немного подрастают за счет запасов желтка.

В конце концов лягушка-нянька добирается до подходящего водоема. Там слизистые "оковы" более или менее долго размокают и наконец растворяются, а освободившиеся головастики получают возможность плавать. Многие дендробатиды приносят своих головастиков к небольшим ручьям; некоторые виды выпускают их в лужицы дождевой воды, скапливающейся в розетках бромелиевых или в пазухах листьев других тропических растений. Лягушки вида *Dendrobates auratus* нередко выпускают своих головастиков в воду, собравшуюся в дуплах древесных стволов. В такой бедной пищей среде головастики *D. auratus* склонны к каннибализму, так что порой лишь один из них доживает до половой зрелости.

По-видимому, в дождевой воде, скапливающейся в пазухах листьев бромелиевых и растений других видов, также мало питательных веществ; это объясняет удивительное явление, которое наблюдал П. Вейгольдт (Фрейбургский университет, ФРГ). Изучая в лаборатории экземпляры *Dendrobates pumilio* (головастики этого вида развиваются именно в таких местообитани-



АРЕАЛ ПЯТИ ВИДОВ дендробатид рода *Phyllobates* имеет дизъюнктивный характер (так называемое пятнистое распределение). Три вида из западной Колумбии — единственные лягушки, о которых известно, что их используют индейцы-охотники при изготовлении отравленных игл для духовых трубок. В Коста-Рике и Панаме встречаются два вида, у которых те же токсины, что и у их колумбийских сородичей, только в несколько меньшем количестве; однако они не распространены там, где индейцы охотятся с духовыми трубками.



ЗАБОТА О ПОТОМСТВЕ, которую проявляет самец или самка, включает и уход за икринками. На снимке — самец вида *Dendrobates silverstonei* охраняет кладку на листе лесной подстилки.

ях), Вейгольдт заметил, что самки регулярно навещают свое потомство и откладывают в воду, где плавают головастики, неоплодотворенные икринки. Значит, головастики питаются икринками.

Иногда биологи пытаются определить для вида относительную "стоимость" родительской заботы о потомстве, ведь такая забота может накладывать на размножение определенные ограничения и даже подвергать родителей дополнительному риску. Компенсируется ли это выгодами, которые вид получает, обеспечивая успешное развитие одного или нескольких выводков отдельной особи? У лягушек-дендробатид родительская забота "дороже", вероятно, в случае видов, у которых роль няньки играет самка, а не самец (например, *D. pumilio*). Самец пары *D. tricolor* из Американского музея мог ухаживать за несколькими кладками икринок одновременно. Более того, при этом он часто издавал призывные крики, видимо, демонстрируя готовность спариться с другими самками, если бы таковые оказались в наличии. Он поглощал пищу, даже нося на спине полный выводок головастиков, которых выпускал в воду уже через несколько часов после вылупления.





ЛЯГУШКИ-"НЯНЬКИ" переносят вылупившихся из икринок головастиков к подходящему водоему. *Слева*: самка панамского вида *Colostethus inguinalis*, на спине у нее 27 головастиков. *Справа*: самец мелкого перуанского вида *Dendrobates reticulatus* несет на спине только одного головастика. Головастик может так сидеть на спине у родителя несколько часов, а может и больше недели.

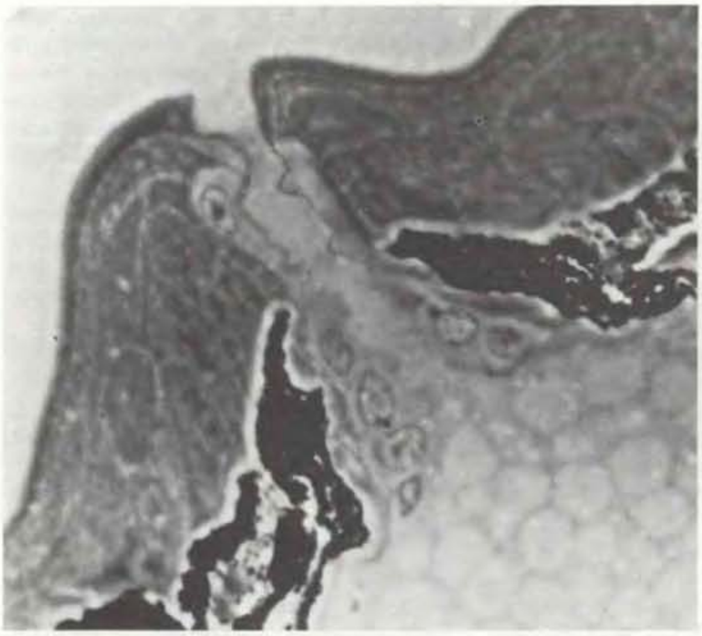
Предполагают, что пол лягушки-няньки (в большинстве случаев самец, иногда — самка) является специфическим для вида признаком. Однако у некоторых дендробатид головастиков переносят к воде как самцы, так и самки. Это трудно объяснить. Чтó в каждом конкретном случае определяет проявление столь сложных поведенческих реакций у того или иного из родителей? Эксперименты с *D. tricolor* наводят на мысль, что обязанности распределяются в результате межполовой конкуренции за роль няньки. В паре *D. tricolor* самка, как и самец, была способна полностью взять на себя заботу о потомстве и проявляла к этому стремление, однако агрессивно охранявший свою территорию самец не давал ей такой возможности. Самка оставалась возле кладки икринок всего час или около того, после чего ей, по-видимому, не разрешалось ее посещать. Однако если самца удаляли из вольеры сразу после оплодотворения, самка ухаживала за икринками и носила на себе головастиков так же успешно, как и он. Если самца удаляли через несколько дней после откладки икринок, а до этого давали ему возможность проявлять родительский инстинкт, самка поедала собственную икру и головастиков. Такого рода каннибализм, возможно, служит у некоторых видов дендробатид механизмом, регулирующим плотность популяции.

Есть еще много неясного не только в естественной истории дендробатид, но и в их эволюционных отношениях и классификации. Как уже говорилось, в это семейство входит приблизительно 130 видов, их разделяют в настоящее время на 4 рода: *Atopophrynus*, *Colostethus*, *Dendrobates* и *Phyllobates*. К роду *Atopophrynus* относится только один вид, к роду *Colostethus* — более 70 видов, главным образом коричневой окраски. Эти виды, за редким исключением, неядовиты. Более 50 видов рода *Dendrobates* и 5 видов рода *Phyllobates* отличаются яркой предупреждающей окраской. Их бросающаяся в глаза расцветка "объявляет" о ядовитых или по меньшей мере неприятных кожных выделениях, весьма эффективно отпугивающих многих потенциальных врагов. Тем не менее ядовитые дендробатиды вовсе не застрахованы от нападения: например, на них успешно охотятся крупные пауки и некоторые змеи. Когда лягушка находится в состоянии стресса, микроскопические железы в ее коже выделяют яд. Хищник, схвативший лягушку, ощущает в пасти сильное жжение или онемение и бросает добычу.

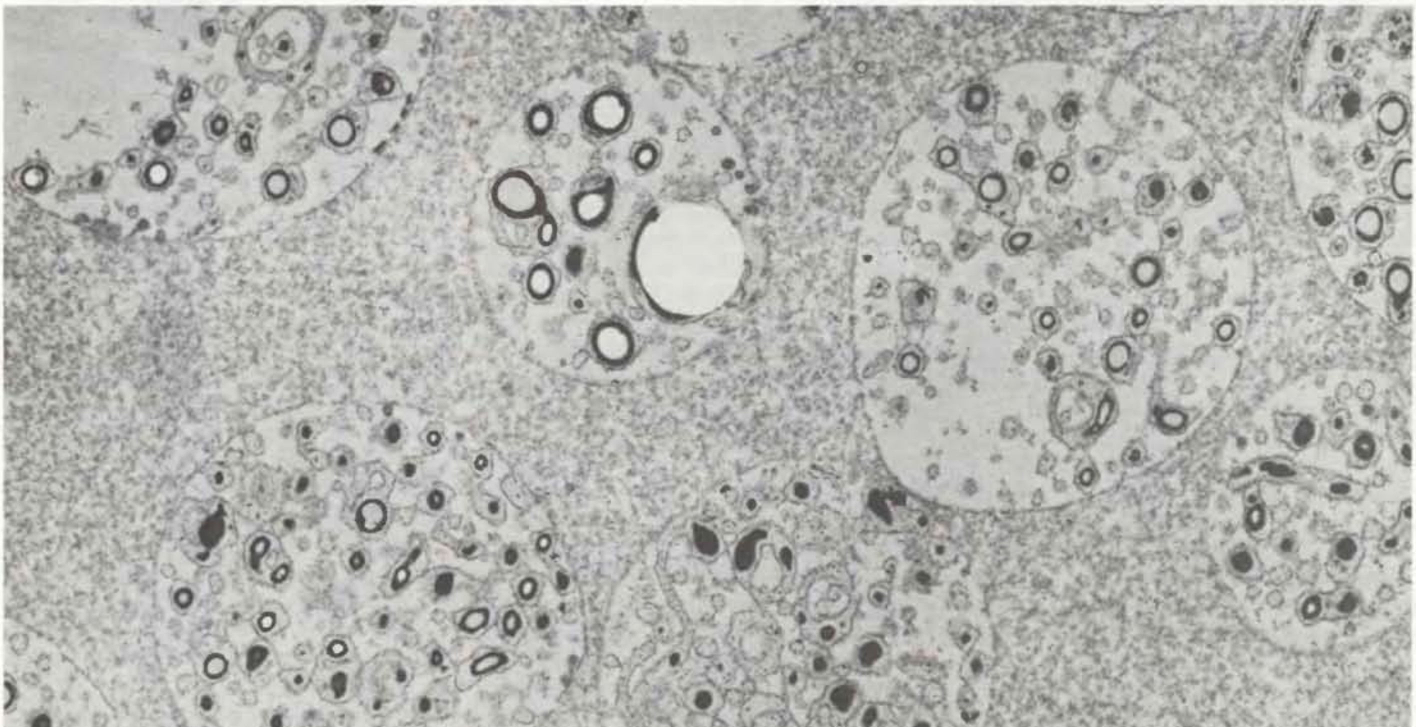
Наиболее ядовиты три вида рода *Phyllobates*, обитающие на западных склонах колумбийских Анд. Эти лягушки выделяют яд, который гораздо сильнее кураре. Кураре популярен к

востоку от Анд, а индейцы западной Колумбии смазывают свои охотничьи иголки ядом *Phyllobates.*

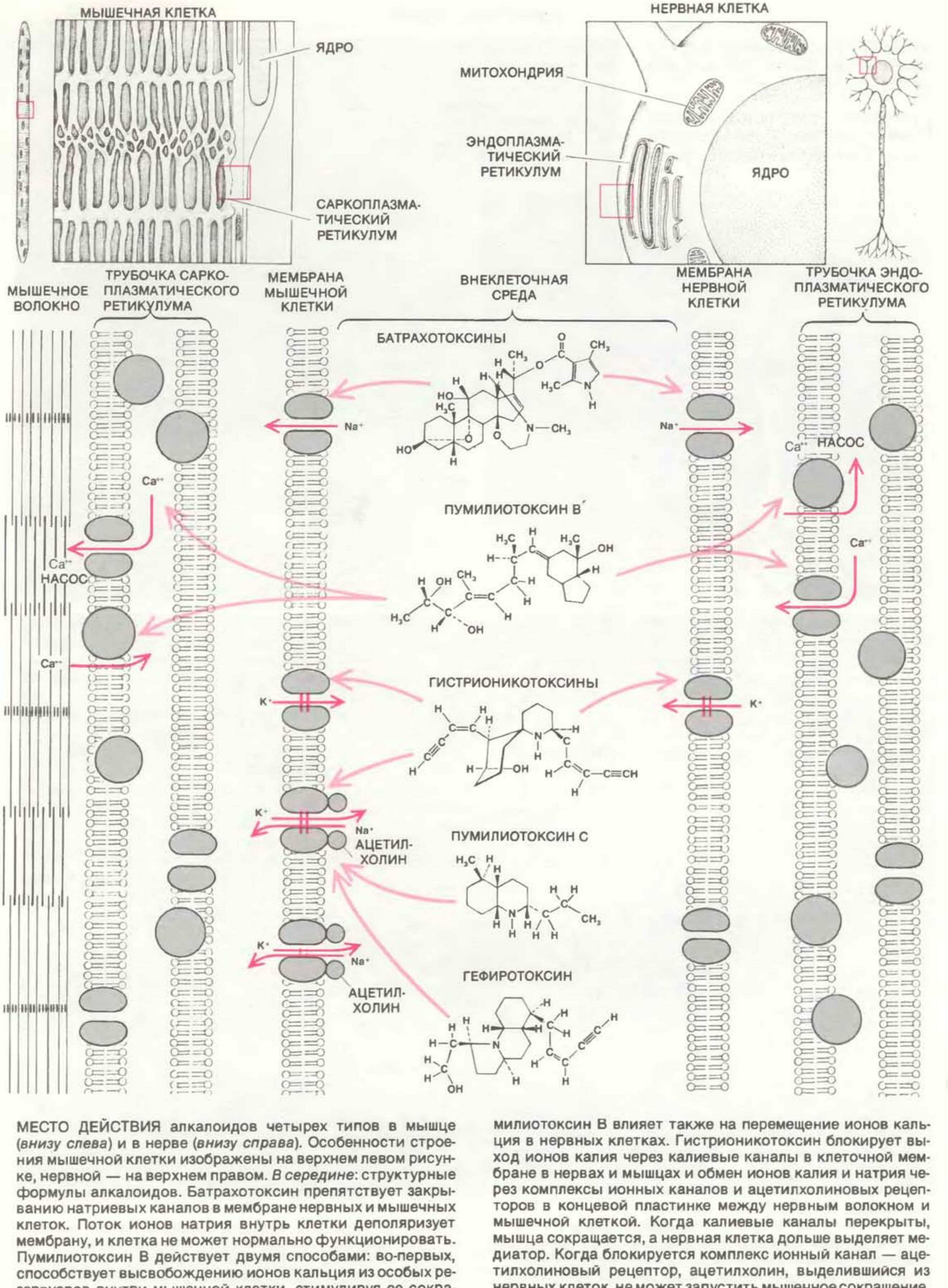
ВЕРОЯТНО, самое раннее упоминание о таком применении лягушек — записки капитана Британского флота Чарлза Стюарта Кочрейна, который путешествовал по Колумбии во время своего отпуска в 1823 — 1824 гг. Пересекая пешком западные Анды, Кочрейн обратил внимание на лягушек, «которых называли rana de veneno ("ядовитая лягушка" по-испански) — длиной около трех дюймов, с желтой спинкой и очень крупными черными глазами». Он записал: «Те, кто использует [их] яд, ловят лягушек на деревьях и сажают их в пустое ведро, и кормят, пока не понадобится яд. Тогда они берут одно из этих







ГРАНУЛЯРНЫЕ ЖЕЛЕЗЫ кожи лягушек-дендробатид. В них накапливаются и, по-видимому, продуцируются токсичные выделения. *Вверху слева*: микрофотография продольного среза через верхнюю часть такой железы (увеличение ×900). Железа заполнена крупными округлыми гранулами. Проток ее проходит через слой пигментных клеток и через эпидермис и открывается на поверхности кожи. *Вверху справа*: электронная микрофотография среза железы (×1500). Пузырьки, или везикулы, в гранулах, вероятно, содержат ферменты синтеза токсинов. Внизу: электронная микрофотография гранул и везикул (×13000). Вверху слева микрофотография железы лягушки вида *Dendrobates tricolor*, остальные — вида *D. auratus*.



МЕСТО ДЕЙСТВИЯ алкалоидов четырех типов в мышце (*внизу слева*) и в нерве (*внизу справа*). Особенности строения мышечной клетки изображены на верхнем левом рисунке, нервной — на верхнем правом. *В середине*: структурные формулы алкалоидов. Батрахотоксин препятствует закрыванию натриевых каналов в мембране нервных и мышечных клеток. Поток ионов натрия внутрь клетки деполяризует мембрану, и клетка не может нормально функционировать. Пумилиотоксин B действует двумя способами: во-первых, способствует высвобождению ионов кальция из особых резервуаров внутри мышечной клетки, стимулируя ее сокращение; во-вторых, подавляет возвращение ионов кальция в эти резервуары, увеличивая длительность сокращения. Пумилиотоксин B влияет также на перемещение ионов кальция в нервных клетках. Гистрионикотоксин блокирует выход ионов калия через калиевые каналы в клеточной мембране в нервах и мышцах и обмен ионов калия и натрия через комплексы ионных каналов и ацетилхолиновых рецепторов в концевой пластинке между нервным волокном и мышечной клеткой. Когда калиевые каналы перекрыты, мышца сокращается, а нервная клетка дольше выделяет медиатор. Когда блокируется комплекс ионный канал — ацетилхолиновый рецептор, ацетилхолин, выделившийся из нервных клеток, не может запустить мышечное сокращение. Пумилиотоксин C и гефиротоксин блокируют ток ионов через комплекс ионный канал — ацетилхолиновый рецептор.

несчастных созданий и просовывают заостренную палку ему в глотку так, чтобы она вышла через одну из задних лап. Такая пытка заставляет бедную лягушку очень сильно потеть, особенно на спине, которая покрывается белой пеной — это и есть самый сильный яд, какой она может дать; в него они окунают концы стрел, и стрелы будут ядовиты целый год. Под белой пеной затем появляется желтое масло, его тоже аккуратно соскребают. Это масло сохраняет свою смертоносную силу 4 или 6 месяцев в зависимости от того, насколько хорошей (так они говорят) была лягушка. Таким способом от одной лягушки получают столько яда, что им можно обработать примерно 50 стрел».

«Путешествия по Колумбии» Кочрейна, наверное, мало кто читал, так что более поздние сообщения о таком любопытном использовании лягушек можно считать независимым подтверждением. Пытались сделать химический анализ образцов яда, соскобленного с индейских охотничьих иголок, но безуспешно. В 1934 и 1955 гг. в Колумбии работал шведский антрополог С. Вассен. Он установил, что ядом лягушек пользовались индейцы чоко двух родственных групп, живших вдоль реки Сан-Хуан. Лягушек, которых он взял у индейцев, зоологи ошибочно определили как *Dendrobates tinctorius*, на самом деле они относились к видам *Phyllobates aurotaenia* и *P. bicolor*. Вассен сумел также достать несколько отравленных иголок у индейцев чоко, которые жили изолированной группой гораздо южнее — на реке Сайха. В этой местности в начале 70-х годов работали и мы вместе с нашим коллегой Б. Малкиным. Собранные нами экземпляры относились к неописанному прежде виду; оказалось, что в них по крайней мере в 20 раз больше яда, чем в лягушках с реки Сан-Хуан. Учитывая необыкновенную токсичность этих лягушек, мы решили назвать этот вид *Phyllobates terribilis*.

Как можно судить по описанию Кочрейна, северные индейцы чоко отравляли свои охотничьи иголки выделениями *P. bicolor* и *P. aurotaenia* еще 150 лет назад. Лягушек насаживали на палку, а иногда еще и держали около огня; затем окунали в их яд оружие. Южные индейцы чоко отравляли охотничьи иголки, просто обтирая их о спину живых экземпляров *P. terribilis*.

В настоящее время необходимость в духовых трубках у индейцев чоко меньше, поскольку стало доступным огнестрельное оружие. Вопреки некоторым встречающимся в литературе описаниям яд дендробатид не употребляется для изготовления отравленных стрел, а использование его в охотничьих иголках, по-видимому, не так уж широко распространено. Это понятно, если учесть, что три наиболее ядовитых вида этого семейства водятся только в довольно небольшой области в западной Колумбии. Два других вида рода *Phyllobates* обнаружены в Центральной Америке, однако они гораздо менее ядовиты, чем их колумбийские сородичи. Кроме того, в их ареал не входит тот район, где индейцы пользуются отравленными охотничьими иголками.

Синтезирующие токсин кожные органы дендробатид представлены так называемыми гранулярными железами, которые открываются микроскопическими порами на поверхности кожи. Ядовитые железы, как и слизистые, разбросаны по всему телу. Гранулярные железы — по-видимому, примитивный признак, общий для всех лягушек; у некоторых видов на его основе развилось защитное приспособление: гранулярные железы стали продуцировать, накапливать и выделять различные биологически активные вещества.

Действующим началом кожных выделений лягушек-дендробатид являются алкалоиды — азотсодержащие циклические соединения. Такие соединения встречаются чаще у растений, чем у животных; однако лягушки-дендробатиды представляют собой богатый источник уникальных алкалоидов. Мы собирали представителей этого семейства по всей области их распространения в тропической Америке и затем проанализировали их защитные выделения. Мы нашли дюжину новых видов ядовитых лягушек и обнаружили более 200 новых алкалоидов, относящихся по крайней мере к пяти различным классам соединений. Химические исследования алкалоидов лягушек были начаты в Национальном институте по изучению артрита, нарушений обмена веществ и заболеваний пищеварительного тракта Б. Уиткопом в начале 60-х годов. За последние два десятилетия основной вклад в эти исследования внесли работы И. Карл, занимавшейся рентгеноструктурным анализом в Национальной научно-исследовательской лаборатории Военно-морского флота, и Т. Токуямы, химика из Университета г. Осака. Большой интерес представляет механизм действия различных алкалоидов дендробатид. Фармакологические испытания этих соединений впервые предпринял Э. Альбукерк из Медицинской школы Мэрилендского университета.

Почти во всех алкалоидах лягушек-дендробатид в структуру молекулы входит так называемое пиперидиновое кольцо, состоящее из одного атома азота и пяти атомов углерода. Простые пиперидиновые соединения обнаружены у всех подгрупп родов *Dendrobates* и *Phyllobates*. Это наводит на мысль, что с эволюционной точки зрения эти два рода — особая ветвь семейства дендробатид. Однако у пяти видов рода *Phyllobates* биосинтез пиперидиновых соединений в значительной степени вытеснен образованием других очень ядовитых алкалоидов — батрахотоксинов (от греческого *батрахос* — лягушка). Батрахотоксины представляют собой основной компонент яда, которым индейцы чоко пропитывают свои охотничьи иглы. Эти сложные алкалоиды родственны стероидам, однако обладают некоторыми особыми свойствами, неизвестными прежде у природных соединений.

Батрахотоксины относятся к наиболее сильным из встречающихся в природе небелковых ядов. Они избирательно повышают проницаемость мембран нервных и мышечных клеток для ионов натрия. В результате развивается необратимая электрическая деполяризация клеток, что вызывает аритмию, фибрилляцию и остановку сердца. По-видимому, батрахотоксины специфически воздействуют на каналы, которые регулируют поток ионов натрия через клеточную мембрану. В нервных и мышечных клетках натриевые каналы играют ключевую роль в проведении электрических импульсов. Связывание батрахотоксинов мешает этим каналам нормально закрываться, в результате чего внутрь клетки устремляется поток ионов натрия и клетка деполяризуется. Нервные клетки больше не проводят импульсы, и мышечные клетки остаются в возбужденном состоянии, в состоянии сокращения. Использование батрахотоксинов в биохимических исследованиях дало новые важные сведения о работе натриевых каналов.

Прочие алкалоиды дендробатид проще батрахотоксинов по структуре и гораздо менее ядовиты, однако и они весьма интересны. Многие пиперидиновые алкалоиды действуют на ионный канал так называемой концевой пластинки в месте контакта нервного волокна и мышечной клетки. С этим каналом связан рецептор ацетилхолина — медиатора нервного возбуждения. В результате взаимодействия медиатора с рецептором в клеточной мембране открывается канал, через который могут проходить и ионы натрия и ионы калия. Если возникающий в результате этого электрический импульс достаточно велик, в мембране открываются особые натриевые каналы. В мышечной клетке этот процесс вызывает высвобождение внутриклеточных ионов кальция, что инициирует сокращение мышцы.

Гистрионикотоксины, необычные спиропиперидиновые алкалоиды, впервые выделенные из лягушек вида *Dendrobates histrionicus*, взаимодейст-



вуют с определенными участками комплекса, в который входят канал концевой пластинки и ацетилхолиновый рецептор, и блокируют прохождение ионов. Это может помешать передаче сигналов от нерва к мышце, так что по внешнему проявлению эффект гистрионикотоксинов сходен с действием кураре. Гистрионикотоксины блокируют также прохождение ионов через особые калиевые каналы. Когда эти каналы открыты, ионы калия выходят из клетки, давая нервной или мышечной клетке возможность вернуться в состояние покоя после передачи нервного возбуждения или после мышечного сокращения. Блокировать эти каналы — значит замедлить проведение нервных сигналов и увеличить продолжительность мышечного сокращения.

Пумилиотоксин В, единственный индолизидиновый алкалоид, впервые выделили из лягушек вида *D. pumilio*; он, похоже, влияет на транспорт ионов кальция. В мышечной клетке вследствие электрического импульса ионы кальция высвобождаются из особых внутриклеточных резервуаров и взаимодействуют с белками мышечного волокна, которое в результате сокращается. По-видимому, пумилиотоксин В не блокирует, а наоборот, облегчает выход кальция и, кроме того, задерживает его возвращение в резервуары. В результате сила мышечного сокращения и его длительность возрастают и в сердечной, и в скелетной мускулатуре. Такой тонизирующий эффект интересен с точки зрения применения алкалоидов этого типа в медицине.

Как ядовитые, так и неядовитые дендробатиды обычно ярко окрашены. Потенциальная опасность представителей этого семейства — причина мимикрии: некоторые безвредные лягушки внешне похожи на ядовитых. Примером может служить вид *Eleutherodactylus gaigeae*: он встречается там же, где *Phyllobates lugubris* и *P. aurotaenia*, а также на территории, лежащей между ареалами двух этих видов, где, вероятно, когда-то обитала их предковая форма. В отличие от ядовитых дендробатид, окраске которых они подражают, неядовитые представители *E. gaigeae* активны ночью. Днем их иногда находят в лесной подстилке, где скрываются от преследователей и виды рода *Phyllobates*. Хотя циклы активности этих животных различны, им, видимо, приходится избегать одних и тех же дневных хищников.

ВОСЕМЬ ЛЯГУШЕК из разных популяций вида *Dendrobates pumilio.* Эти экземпляры — пример сильной изменчивости, которая наблюдается среди лягушек комплекса популяций *D. pumilio* карибского побережья западной Панамы. На рисунке все лягушки одинаковы по размеру, однако на самом деле величина их варьирует не меньше, чем расцветка. Кроме того, различаются и поведение, и состав ядовитых выделений.

У *Phyllobates* отмечается меньшая изменчивость, чем у *Dendrobates*, которые удивительно разнообразны по окраске и узорам на коже. Многие представители рода *Dendrobates* просто полосатые, что, по-видимому, следует считать более примитивным по сравнению со сложной пятнистой расцветкой других дендробатид. Изменчивость пятен, крапинок, полосок может наблюдаться даже между популяциями одного и того же вида. Интересно, каковы причины такого разнообразия и не связано ли оно с образованием новых видов? Не исключено, что ответы на эти принципиальные вопросы даст изучение одного из центральноамериканских видов — *D. pumilio*. *D. pumilio* — мелкие лягушки; обычно они встречаются не выше 500 м над уровнем моря, в дождевых лесах карибского побережья Никарагуа, Коста-Рики и западной части Панамы. В Никарагуа и Коста-Рике изменчивость этого вида относительно невелика; здесь лягушки, как правило, красные или красновато-оранжевые с черными или светло-голубыми задними лапами. В Панаме же вид *D. pumilio* необычайно изменчив. Это связано с геологической историей западной Панамы. Около 12 000 лет назад уровень моря начал подниматься, вследствие чего возник архипелаг прибрежных островов, а на прилежащей части материка выросли дождевые и низинные заболоченные леса. За прошедшее с тех пор время архипелаг отделялся от материка и вновь соединялся с ним, по-видимому, несколько раз. В этом регионе встречаются популяции лягушек *D. pumilio* всех цветов радуги от красного до синего, а в одной из популяций лягушки имеют совсем необычную окраску — они черно-белые. Разные популяции выглядят как разные виды; мы полагаем, что в действительности следует говорить о по меньшей мере двух самостоятельных видах. Тем не менее большинство популяций принадлежит к одному виду — собственно *D. pumilio*.

Эта изменчивость обусловлена не просто независимой эволюцией изолированных островных популяций. На одном и том же острове могут жить сильно различающиеся популяции. Известны и примеры значительной внутрипопуляционной изменчивости. Один из таких случаев мы обнаружили на северо-западной окраине острова Бастиментос и изучали его несколько лет. Окраска спины у здешних лягушек изменяется от оранжево-красной через бронзовую до металлически-зеленой. Среди экземпляров, которых мы ежегодно отлавливали в прибрежном лесу, оранжево-красные лягушки, как правило, составляли около 96%, и только 4% имели другую окраску. Однако дальше от берега, в тенистых зарослях какао, оранжево-красных лягушек было уже только 84% и даже меньше, а доля бронзовых и зеленых лягушек доходила до 32%. Как видим, популяции одного и того же вида, заселяющие различные местообитания, отличаются друг от друга — это пример так называемого сбалансированного полиморфизма, который, вероятно, является генетической основой некоторых резких различий, наблюдаемых среди представителей вида *D. pumilio*.

Хотя территории сильно отличающихся друг от друга популяций *D. pumilio* обычно разделены участками леса, в которых лягушки этого вида не встречаются, в континентальной части Панамы мы обнаружили область, где жили практически однородная популяция красных лягушек и такая же чистая популяция их зеленых сородичей. В обеих популяциях лягушки другого цвета встречались крайне редко, но иногда попадались особи совершенно иной окраски, напоминавшие лягушек из более отдаленных районов. Эти данные лишний раз говорят о том, что различающиеся по окраске популяции лягушек могут возникать на основе одного общего генофонда.

У ПРЕДСТАВИТЕЛЕЙ *D. pumilio* варьирует не только расцветка. Популяции этих лягушек различаются также по размерам и поведению особей. Известны наземные и древесные формы; в присутствии потенциальных врагов одни ведут себя скрытно и осторожно, другие держатся с демонстративным бесстрашием. Ни один из этих признаков, по-видимому, не коррелирует с интенсивностью кожных защитных выделений (она, кстати, также сильно варьирует). В популяциях *D. pumilio* западной Панамы встречаются в общей сложности более 80 различных алкалоидов, среди них можно найти представителей всех основных классов пиперидиновых токсинов. Число различных алкалоидов у лягушек в одной отдельной популяции колеблется от 6 до 24, причем многие из них свойственны только лягушкам этой популяции и в других популяциях не обнаруживаются. По-видимому, *D. pumilio* можно считать самым изменчивым видом среди позвоночных. Правда, почему это так, пока что совершенно неясно.



Веерные своды

В XIV в. английские зодчие преобразовали готическую систему остроконечных арок в свод — оболочку из плавно сопряженных поверхностей. Конструкция эта отличается не только изяществом — главное, к чему стремились ее создатели, — но и высокой прочностью

УОЛТЕР К. ЛИДИ, мл.

ПРОБЛЕМА перекрытия внутренних пространств всегда была одной из основных в архитектуре. Сложность задачи возрастает по мере увеличения пролета между опорами, и естественно, что особое значение она приобрела в строительстве средневековых соборов и монастырей. При их высоте зрительное движение стен навстречу друг другу фокусирует внимание на перекрытии, пространственный рисунок которого призван поражать воображение. Проблема осложнялась тем, что перекрытие, как правило, выполнялось из камня и было чрезвычайно тяжелым*. Решение, использованное при строительстве большинства известных готических храмов, сводилось к тому, чтобы возводить перекрытие в виде последовательности сводов, расчлененных на почти плоские выложенные из камня поверхности, заключенные в раму из остроконечных арок. Арки играли роль ребер, принимающих на себя по крайней мере часть нагрузки.

По мере накопления средневековыми мастерами опыта возведения ребристых сводов они пришли к выводу, что конструкция может стать легче, а ее рисунок свободнее, если ввести в структуру перекрытия еще и дополнительные ребра. Постепенно сводчатая конструкция превратилась в сложный кружевной рисунок из ребер-нервюр, выполняемый с бесконечным разнообразием. В Англии в XIV в. умножение числа ребер привело к изобретению совершенно новой конструкции — веерного свода. В этой конструкции слабо выпуклые каменные поверхности ребристого свода уступают место коноидальным поверхностям, напоминающим половину конуса, поставленную на острие. Для перекрытия каждого пролета коноиды соединялись развернутыми вверх основаниями по четыре. Многочисленные ребра расходились вверх вдоль поверхности каждого коноида — именно этому веерному рисунку из нервюр новый вид свода обязан своим названием. Промежутки между ребрами покрывались сложной каменной резьбой.

В ребристом своде всю нагрузку принимают на себя арки. В веерном своде коноид ведет себя как оболочка двоякой кривизны, в которой напряжения относительно равномерно распределены в слое камня. В то время строители еще не умели анализировать распределение усилий в своде-оболочке, тем не менее созданный ими веерный свод оказался очень устойчивой конструкцией. За период между 1350 и 1540 г. в Англии было возведено более сотни таких сводов, размеры которых постепенно увеличивались. Среди наиболее знаменитых — перекрытие капеллы Кингз Колледж в Кембриджском университете и капеллы Генриха VII в Вестминстерском аббатстве.

Веерные своды храмов поражают своим великолепием. Вдоль стен коноиды перекрытия зрительно переходят один в другой, создавая впечатление мягко вылепленного пространства в противоположность угловатости рисунка обычного ребристого свода. Тонкая резьба на поверхности коноида подчеркивает его форму, усиливая ощущение зрительной целостности зала. По всей видимости, именно этого визуального эффекта желали достичь создатели веерного свода. Не следует забывать, однако, что стремление к изяществу конструкции привело их к решению фундаментальной строительной задачи, несмотря на то что из-за отсутствия теоретических познаний им приходилось опираться только на собственный опыт.

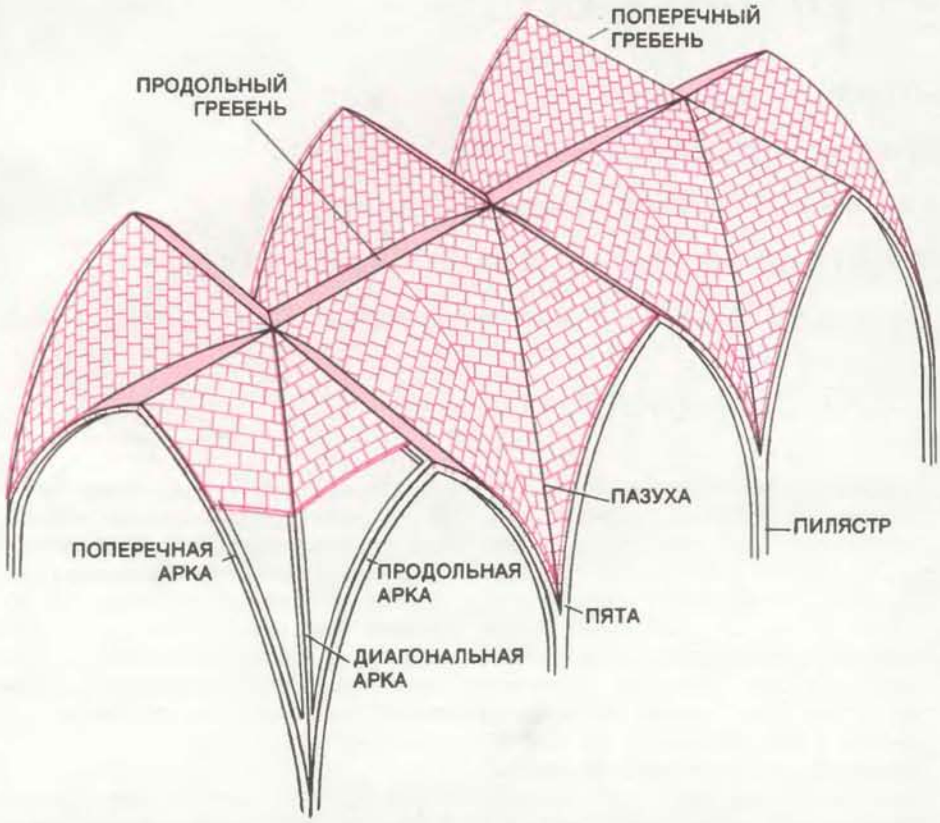
ЭВОЛЮЦИЮ веерного свода определяли умелое использование физических свойств камня и художественный вкус английских зодчих и их заказчиков. Камень выдерживает значительное сжатие, но имеет малую прочность на растяжение. За счет большой массы в горизонтальной балке, поддерживаемой с двух сторон, возникают значительные растягивающие усилия. В каменном перекрытии такие усилия должны быть преобразованы в усилия сжатия. Так, в остроконечной арке ребристого свода усилия, направленные вниз, преобразуются в усилия сжатия, которые передаются через каменную кладку арки стенам.

Единицей или модулем крестового свода является прямоугольный в плане пролет, продольная ось которого перпендикулярна к главной оси зала. В ребристом своде пролет образуется тремя парами арок. Поперечные арки перекрывают зал, продольные — пролеты в плоскости стен. Под продольными арками зачастую располагался верхний ряд больших окон для освещения хоров. Противолежащие углы каждого пролета связываются диагональными арками, пересекающимися в центре модуля. Высокие столбы — пилястры, подведенные под пяты — точки, где встречаются нижние сегменты всех трех арок, поддерживают весь свод на большой высоте над уровнем пола. Вес перекрытия передается фундаменту через столбы, а горизонтальный распор —приставным столбам-контрфорсам посредством наклонных висячих арок (аркбутанов), отходящих от стены примерно на уровне пят арок.

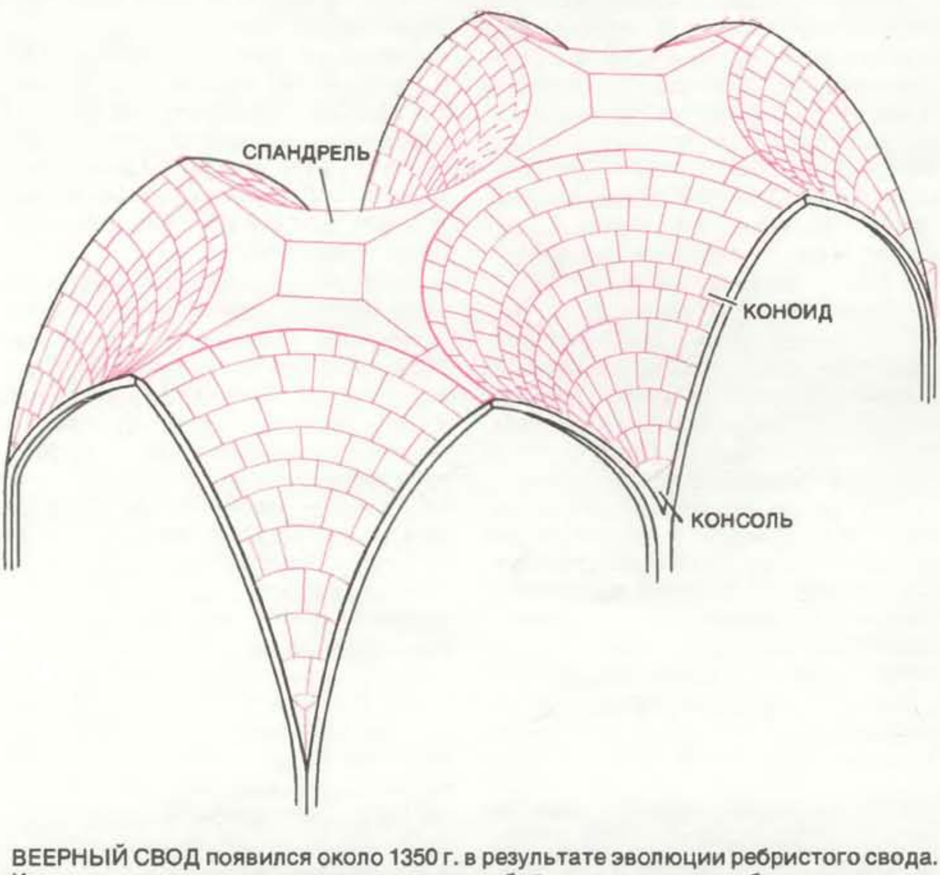
При сооружении ребристого свода в первую очередь возводились все арки с

* В Англии, где развитие корабельного искусства повлекло за собой совершенствование мастерства плотников и резчиков по дереву, возводилось множество деревянных перекрытий. Среди них немало имитаций каменного веерного свода (особенно это характерно для XVI в.). Наиболее знамениты из них Большой зал и капелла Генриха VIII во дворце Хэмптон Корт, близ Лондона. — *Прим. перев.*

ВЕЕРНЫЕ СВОДЫ КАПЕЛЛЫ КИНГЗ КОЛЛЕДЖ в Кембриджском университете — одни из самых красивых. Возведение изображенного здесь главного свода капеллы завершилось в 1515 г. Это веерный свод с самым большим пролетом — около 12,7 м. Взаимопересекающиеся коноидальные поверхности воспринимают вертикальную нагрузку перекрытия. Называемые коноидами, они ведут себя как оболочки, в которых все напряжения распределены относительно равномерно. Рисунок резьбы, созданный по образцу рисунка витражных окон, объединяет интерьер в единое целое.



РЕБРИСТЫЙ СВОД, классический пример решения перекрытий в готике, трактовался средневековыми мастерами как каркасная конструкция, в которой вес перекрытия принимают на себя ребра стреловидных арок. Каждый модульный элемент перекрытия (пролет) образуется шестью арками, собранными в прямоугольную по плану ячейку с двумя пересекающимися диагоналями. Поверхность между ребрами заполняется плоскими каменными плитами.



ВЕЕРНЫЙ СВОД появился около 1350 г. в результате эволюции ребристого свода. Коноиды веерного свода представляют собой тела вращения, образуемые поворотом кривой вокруг вертикальной оси. По отношению к центральной точке пролета горизонтальное сечение коноида — выпуклая кривая, а вертикальное — вогнутая, поэтому коноид ведет себя как система вертикальных и горизонтальных арок, передающих нагрузку стене.

помощью сложной системы деревянных лесов, поддерживавших кружала, по которым производилась кладка рядов камня. Шесть арок каждого пролета служили каркасом для выкладывания ребристого свода.

Ребристые своды знаменитых соборов, например в Руане или Амьене, отличаются чрезвычайной четкостью линий, свойственной готическому стилю. Этот эффект достигался всемерным подчеркиванием роли нервюр как обрамлений, заполняемых равномерной каменной кладкой. Кроме того, зрительно выделенные поперечные арки усиливают ритмическую расчлененность перекрытия на отдельные пролеты. Веерные своды создают совершенно иной эффект: образованные также линейными элементами, они подчеркивают, однако, единство интерьера как художественного целого. Взгляд скользит по потолку, увлекаемый вперед мягкой ритмикой коноидальных форм, и разграничение пролетов, как правило, не имеет существенного значения.

Несмотря на контраст стилистических особенностей двух типов сводчатых перекрытий, исторический переход от одного к другому произошел не сразу. Веерный свод появился в процессе постепенной эволюции ребристого. Последний получил распространение как в Англии, так и в странах континентальной Европы, тогда как первый — исключительно в Англии. Интересно проследить причины, почему новая форма сводов встречается только в Англии.

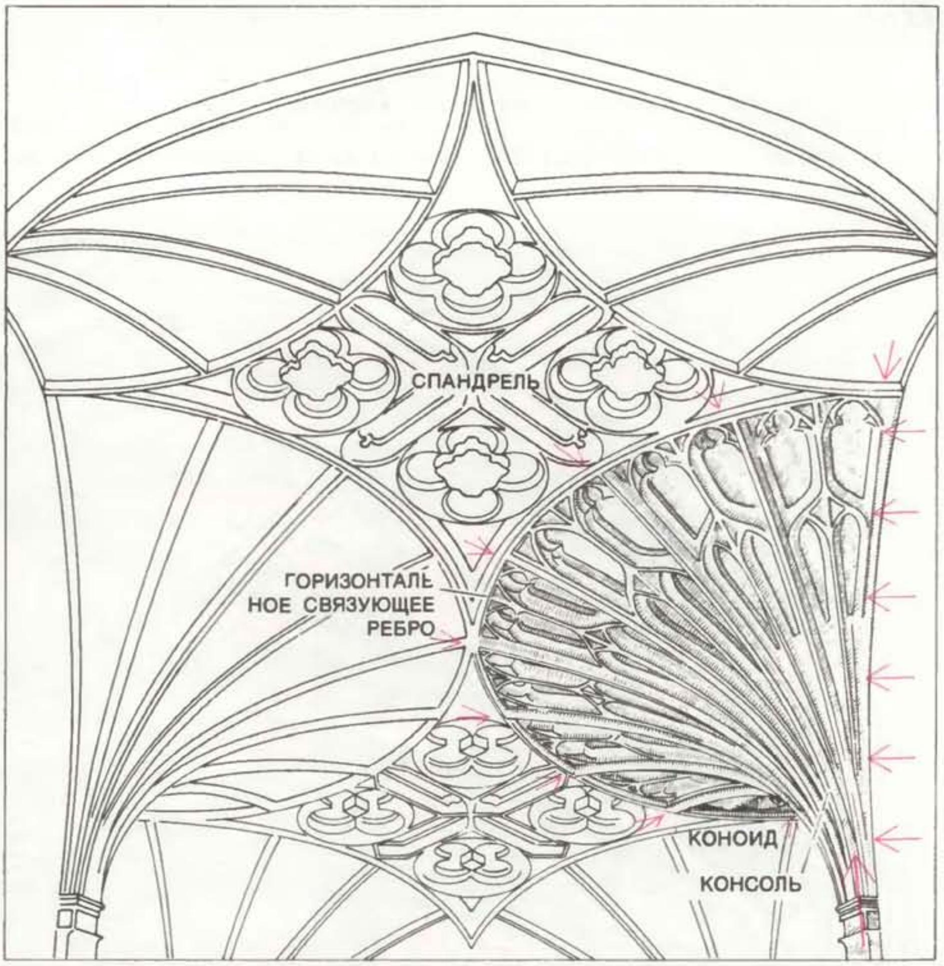
С этой целью обратимся к особенностям художественного стиля и строительной техники средневековых Англии и Франции. Французские мастера производили кладку рядов камня между арками параллельно сторонам перекрываемого пролета, и поэтому свод вырастал вверх горизонтальными слоями. Каждый квадрант пролета подобен остроконечной кровле; гребень квадранта считается продольным или поперечным в зависимости от его ориентации. Вдоль шва, соединяющего соседние квадранты, ряды камня выкладывались параллельно обоим гребням, то есть под прямым углом к сторонам пролета. Сначала английские мастера следовали той же практике, но вскоре они стали выполнять кладку рядов камня от гребня к шву (пазухе) примерно под углом в 45° к обоим гребням. Такой способ кладки уже содержал в себе зародыш округлой структуры, превратившейся впоследствии в коноид.

Имелись и другие расхождения в строительной технике обеих стран. Во Франции кладка каменных рядов свода производилась поверх ребер; ряды эти скреплялись с ребрами только с

помощью раствора. В Англии в верхних частях ребер делались глубокие борозды, в которые вставлялись крайние камни каждого ряда кладки перед заливкой их раствором. Этот прием, известный как шпунтовка, связывал ребро и поверхность свода в единое целое — подтверждение тому, что английские каменщики считали ребро несущим элементом конструкции, а не только поддерживающим.

Даже по своей конструкции английские арки отличаются от французских. Во Франции наибольшее значение придавалось высоте*, освещенности сооружения и обнаженности его архитектурного рисунка. Для увеличения размера окон, освещающих хоры, продольные арки нередко приподнимали, подводя дополнительные тонкие пилястры под их пяты. Такой прием влек за собой существенные изменения конструкции. Поверхность свода, примыкающую к окну, требовалось дополнительно искривлять для согласования с формой арки. К тому же значительно уменьшалась площадь стены между соседними окнами, и распор от свода приходился на узкую полосу стены. Поэтому для аркбутана необходимо было отыскивать оптимальную позицию, в противном случае стена могла обрушиться.

В Англии высоте сооружения и выявленности каркаса придавали меньшее значение. Стены, как правило, были и ниже, и массивнее. Окна для освещения хоров имели меньшие размеры, и потому не требовалось приподнимать продольные арки. Соответственно увеличивалась площадь участка стены, на который приходился распор от свода, и проблема контрфорсов не была столь острой.



КОНОИД ВЕЕРНОГО СВОДА должен быть сжат по всем своим границам, чтобы сохранять устойчивость. На рисунке показан свод храма святой Марии в Нортлейе (Оксфордшир), завершенного около 1440 г. Спандрель — тяжелая плита, расположенная между коноидами, создает необходимое усилие сжатия вдоль горизонтального разделительного ребра и служит замковым камнем для всего свода. Снизу коноид опирается на консольный выступ стены. Массивные стены принимают на себя нагрузку, передаваемую через оболочку коноидов. Резные ребра на поверхности коноидов являются скорее декоративными, чем конструктивными элементами.

Ребристый свод распространился в большинстве европейских стран в течение XII столетия. Вскоре было установлено, что добавление к основным типам ребер, двух других типов придает живость рисунку сводов. Ребро первого типа — «тьерсерон» поднимается от пяты к гребню примерно по биссектрисе угла между диагональным и поперечным или продольным ребрами. Ребро второго типа — «льерн» — представляет собой короткий элемент конструкции, расположенный между основными ребрами.

Английские мастера возводили во множестве своды с добавочным ребром — тьерсероном. Это дополнение соответствовало английской манере, в частности оно позволяло достичь большей плавности линий. В отличие от французской манеры, для которой характерно стремление к высоте и освещенности, нередко за счет уплощения поверхностей, английские зодчие воспользовались тьерсеронами, расходящимися вверх от пяты арки, для создания пространственного рисунка, отличающегося живостью криволинейных очертаний.

В высоких ребристых сводах пространство, обозначаемое пересечением поверхностей вдоль пазухи, утрачивало правильность ближе к углу пролета, поэтому потребовалось введение дополнительных ребер. По мере увеличения их числа это пространство становилось все более правильным, а «паутину» ребер стали рассматривать как самостоятельную художественную тему, достойную дальнейшей разработки. В то же самое время английский способ кладки обнаружил ряд интересных особенностей. Как отмечалось выше, английские мастера выкладывали ряды камней свода под углом к продольному и поперечному гребням, и при введении тьерсеронов ряды кладки в углах пролета стали напоминать правильные горизонтальные многоугольники, концентрически поднимающиеся вверх от пяты арки к центральной точке свода, увеличиваясь в размере.

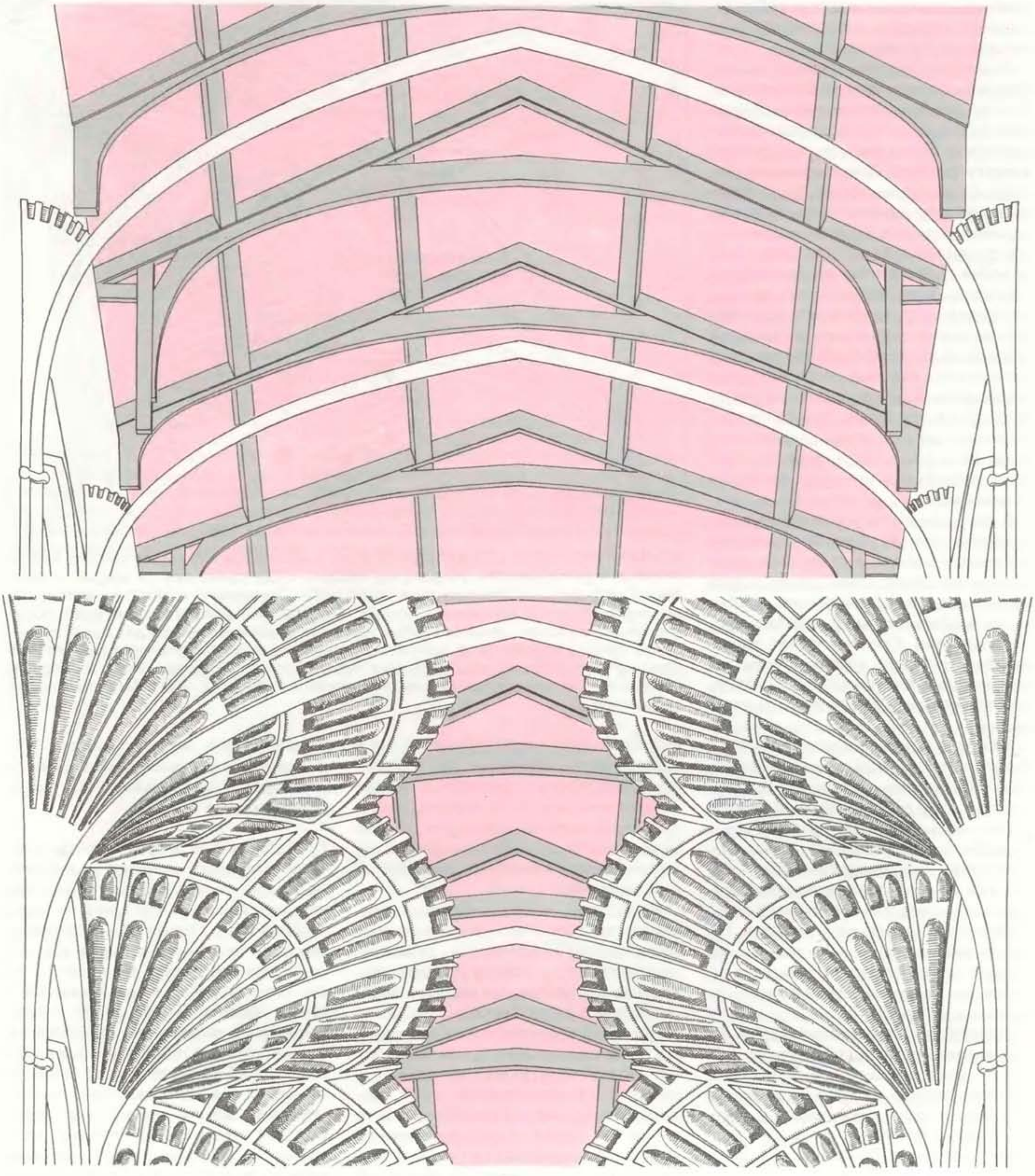
К НАЧАЛУ XIV в. свод, в котором использовались дополнительные ребра, концентрическая кладка и шпунтовка, представлял собой почти готовый веерный свод. Однако для «рождения» веерного свода понадобились еще два нововведения. Первое состояло в том, что сегмент ребра и примыкающий к нему с обеих сторон фрагмент поверхности свода высекаются из одного каменного блока. Отформованные камни соединялись затем примерно таким же способом, как и ледяные блоки при сооружении иглу — эскимосской хижины. Эта строительная техника, основанная на пригонке блоков, получила широкое распространение при возведении веерных сводов.

Применение техники пригонки блоков английскими строителями отнюдь не было мотивировано чисто техническими соображениями. Напротив, есть

* В крупнейших готических соборах Франции, Германии и Испании верхняя точка центрального свода в среднем имеет высоту около сорока метров над уровнем пола. — *Прим. перев.*

все основания полагать, что в немалой степени оно обусловлено стремлением к совершенству форм и рисунка. Английских мастеров привлекала возможность выполнения поверхности свода в виде крупных сочлененных элементов, поскольку добиться чистоты каменной резьбы было, естественно, гораздо легче на относительно крупных каменных блоках, плотно пригнанных один к другому, чем на каменной кладке, заполнявшей пространство между ребрами. По мере увеличения размера блоков они начали играть и конструктивную роль.

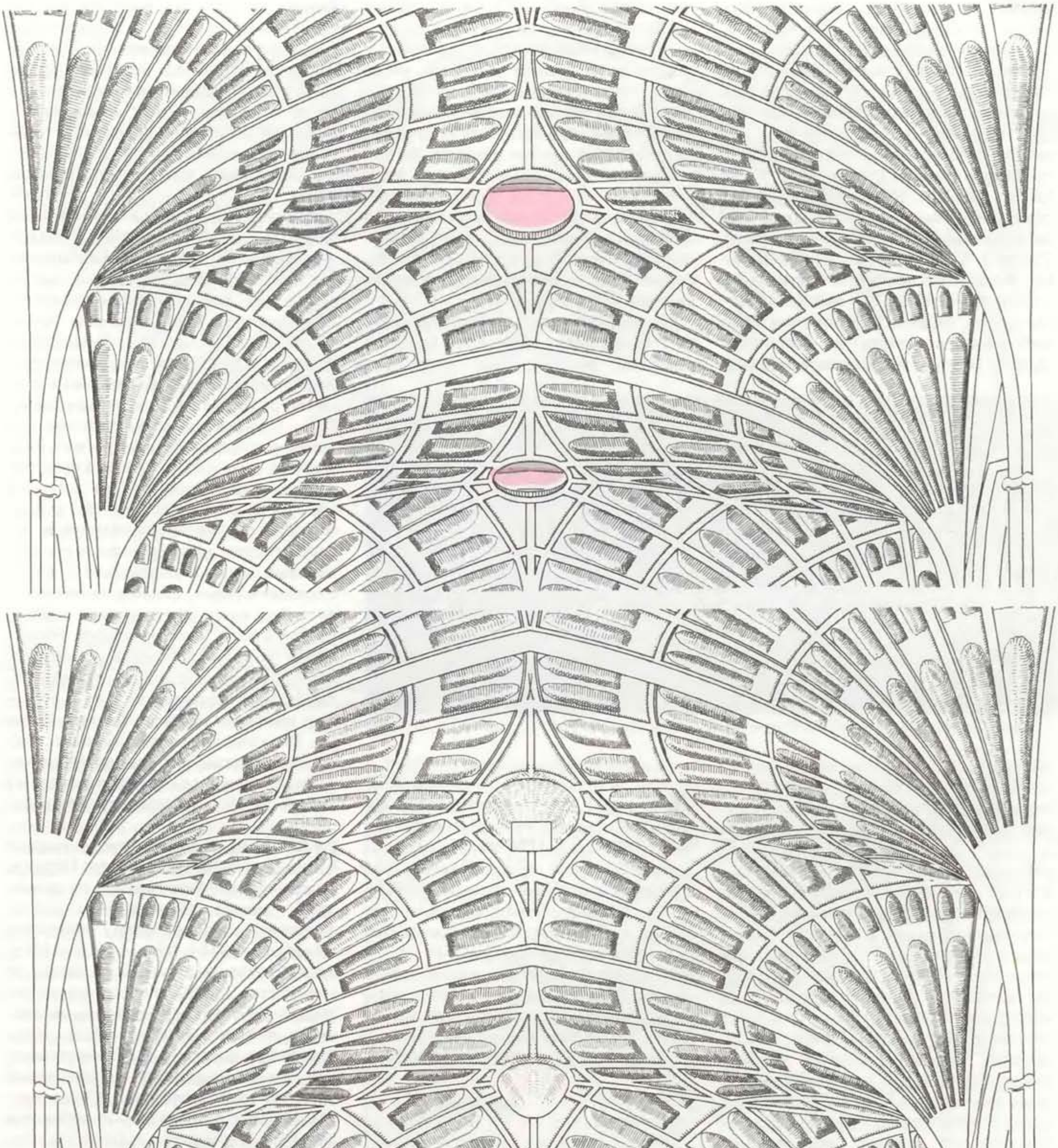
Другое нововведение явилось следствием первого и состояло в следующем. В обычном ребристом своде попереч-



ВОЗВЕДЕНИЕ СВОДОВ в капелле Кингз Колледж осуществлялось в четыре этапа. Каменные блоки обрабатывались в мастерских неподалеку от строительной площадки. Их лицевые поверхности подвергались в этих мастерских лишь частичной обработке. В процессе возведения сводов использовались деревянные леса и кружала. Только после того, как поднимались на полную высоту стены и на них устанавливались конструкции кровли, возводились поперечные арки, разделяющие перекрытие на отдельные пролеты (*слева вверху*). Затем каменщики выводили коноиды, после-

ное сечение ребра, как правило, сохраняет вертикальность. Однако при использовании техники крупноблочной пригонки шов между двумя соседними каменными блоками удается сделать идеально чистым только в том случае, если ось сечения ребра перпендикулярна поверхности свода. Умножение числа ребер, а тем самым и их пересечений, привело к тому, что это нововведение стало играть важную роль в художественном отношении.

К середине XIV в. в Англии уже существовали все главные компоненты веерного свода. По мнению большинства современных исследователей, отличительными особенностями веерного свода являются: наличие образу-



довательно отступая от поперечных арок к центру пролета (*слева внизу*). Все коноиды сооружались одновременной кладкой горизонтальными рядами. Когда горизонтальные арки соседних коноидов смыкались в центре пролета, между ними укладывался крупный камень поперечного гребня, служащий замковым камнем для обеих арок. Затем устанавливались камни продольного гребня (*справа вверху*), и в центр каждого спандреля сверху опускался крупный тяжелый босс (*справа внизу*). После снятия временных кружал выполнялась окончательная отделка свода.

ющих свод коноидов правильной геометрической формы; правильная схема распределения ребер на поверхности свода с соблюдением постоянства их кривизны; четко выделенный спандрель или центральная уплощенная пазуха свода; перпендикулярность сечения ребер поверхности свода; наконец, резная его поверхность. Кажется, весь набор перечисленных характеристик встречается впервые в маленьких, квадратных в плане, резных сводах над гробницами знатных людей. Мы не в состоянии проследить начальную стадию развития веерных сводов, поскольку многие памятники были уничтожены во времена Генриха VIII. Из-за малого размера сводов гробниц их нередко называют декоративными или «игрушечными». Перекрытие капеллы Тринити Чэпел в аббатстве Тюйксбери, сооружение которого завершилось примерно в 1380 г., является, по-видимому, первым применением веерных сводов в архитектуре церковных зданий. Первым же крупноразмерным веерным сводом и одним из самых значительных примеров его ранних форм является свод в монастыре при кафедральном соборе в Глостере, возведение которого было начато во второй половине XIV в.

Важнейшим элементом веерного свода является коноид — тело вращения, образованное поворотом кривой вокруг вертикальной оси. Эта ось проходит через угловую точку пролета, а образующей кривой является ребро свода. Естественно, что любое горизонтальное сечение коноида представляет собой окружность или ее дугу. По отношению к центру пролета окружность является выпуклой кривой, а линия ребра — вогнутой. Таким образом, в любой точке поверхности коноида выпуклая кривая, лежащая в горизонтальной плоскости, пересекается с вогнутой кривой, принадлежащей вертикальной плоскости.

Поскольку коноид, «вырастающий» из угла пролета, образуется концентрическим расхождением вверх дуг окружности, квадратный в плане пролет идеально соответствует веерному своду (такие своды возводились иногда и над прямоугольными модулями зала, но для согласования двух противоречащих друг другу форм требовалась уже изощренная изобретательность). В случае квадратного плана пролета сечением каждого коноида оказывается четверть окружности. В верху свода, в его центре, при пересечении четырех коноидов образуется ромбовидная пазуха. Для ее заполнения и служит спандрель. Граница между коноидами и спандрелем зачастую обводилась горизонтальным ребром для зрительного разграничения двух структурных форм.

Средневековые строители не были формалистами и достаточно свободно видоизменяли те или иные компоненты веерного свода сообразно своему вкусу или необходимости изменить конструкцию. Так, в капелле Кингз Колледж ребра кажутся расположенными на одинаковом расстоянии друг от друга, но в действительности промежутки между ними не совсем одинаковы. Причина заключается в том, что решение построить перекрытие в виде веерных сводов было принято после частичного возведения здания в старой манере. В аббатстве Шерборн в Дорсетшире сохранились прекрасные веерные своды, архитектурному рисунку которых широко подражали в XV в. Горизонтальные сечения коноидов в алтарной части храма имеют, однако, не круглую, а многоугольную форму. Более того, для образования коноидов использовались отдельные ребра и заполняющая каменная кладка, а не метод пригонки крупных блоков. И все же, несмотря на эти расхождения с идеальным типом, оба перекрытия несомненно относятся к семейству веерных сводов по общему характеру решения.

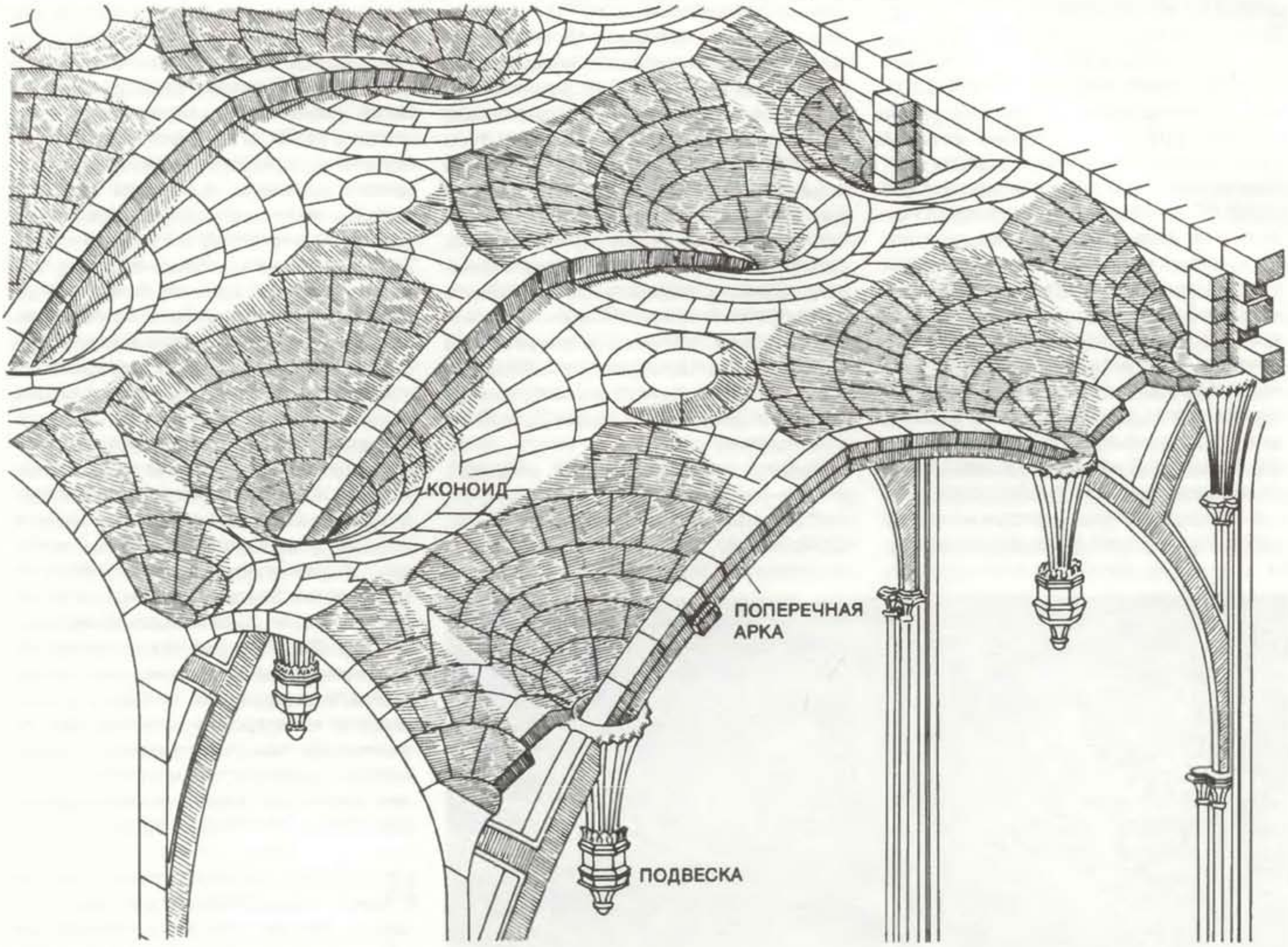




МАРКИРОВКА КАМНЕЙ позволяла строителям капеллы Кингз Колледж собирать потолки, следуя знакам, которые высекались на каменных блоках в мастерских. Кружок в центре каждого камня обозначает квадрант пролета, цифры с краю — номер горизонтального ряда кладки коноида, позицию в этом ряду и очередность установки камня на место. Маркировка камней еще различима на лицевой поверхности свода и выполнена так, как показано на плане квадранта главного свода.

ИСТОРИЮ веерного свода следует разделить на три периода. Первый, включающий сооружение «игрушечных» сводов, завершается строительством алтарной части Глостерского собора в 1412 г. Между 1412 и 1430 г. большепролетные веерные своды не строились. Это было время острой нехватки рабочей силы, непосильных налогов и экономического спада. Веерные своды дороги, и практически некому было выделить достаточные средства на их возведение. Строительство сводов над алтарем в аббатстве Шерборн в конце 30-х годов XV в. знаменует начало второго периода, длившегося до середины 70-х годов. Наконец, в течение третьего периода — с 1475 по 1540 г. — было воздвигнуто наибольшее количество самых больших и самых интересных перекрытий из веерных сводов.

Абсолютное большинство веерных



СВОДЫ КАПЕЛЛЫ ГЕНРИХА VII в Вестминстерском аббатстве, несомненно, самые роскошные. Их возведение началось примерно в 1500 г. Уникальность конструктивного решения перекрытий заключается в комбинировании каркасов и оболочек. Мощные поперечные арки несут на себе крупные расположенные ближе к стенам подвески, от которых «вырастают» коноиды. Поперечные арки пересекают поверхность свода в месте крепления подвесок и проходят сверху, поэтому они невидимы изнутри в центральной части перекрытия. Основная вертикальная нагрузка приходится на поперечные арки там, где они пересекают свод. Арки являются каркасными элементами, коноиды — оболочками.

сводов возведены в церковных зданиях, главным образом в капеллах, сооружавшихся для отправления заупокойных служб по знатным и богатым завещателям, пожертвовавшим свои вклады. Однако, помимо молитв за упокой души жертвователя, воспевалась и его земная слава, поэтому веерный свод с эффектностью его форм и пышностью резного декора как нельзя лучше соответствовал этому двойному назначению.

Во второй половине XV в. интерес к веерным сводам капелл стали проявлять буржуа и король. Веерные своды капеллы Сейнт Джордж Чэпел в Виндзорском замке датируются 1480 г. Вскоре после 1500 г. началось возведение наиболее роскошных веерных сводов капеллы Генриха VII в Вестминстерском аббатстве, которая по замыслу Генриха VII призвана увековечить память о династии Тюдоров. Таким образом, церковь, знать, буржуа и корона — все внесли свой вклад в распространение новой архитектурной формы.

Распространение веерных сводов сопровождалось увеличением их размеров. Возведенные в 1380 г. своды капеллы Тринити Чэпел в аббатстве Тюйксбери имели всего 1,7 м в пролете. Главные своды капеллы Кингз Колледж, строительство которой завершилось в 1515 г., имели пролет, равный 12,7 м. Это крупнейший из пролетов, и причина, по которой английским мастерам понадобилось полтораста лет, чтобы довести пролет веерного свода до максимальной величины, заключалась в том, что они весьма смутно представляли себе, как ведет себя эта конструкция.

ИНЖЕНЕРЫ XX в., используя современную технику расчетов, показали, что сохранение устойчивости коноида веерного свода может быть достигнуто только в том случае, если он равномерно нагружен по всем граничным контурам. Особое значение имеет нагрузка, приложенная по линии, разделяющей верхний обрез коноида и спандрель. Она создается самим спандрелем — массивной каменной плитой с рельефным резным блоком в центре, называемым «боссом». Босс в капелле Кингз Колледж, например, весит почти полторы тонны.

Спандрель играет роль замкового камня для вертикальных арок. Как и в арке ребристого свода, усилия направлены в стороны и вниз. Вертикальная нагрузка передается основанию коноида в той точке, где он примыкает к стене. Каждый коноид опирается на консоль — выступ стены в углу пролета. Как уже отмечалось, коноид выкладывался концентрическими рядами, расширяющимися при подъеме к центру пролета, поэтому каждый ряд кладки имеет форму горизонтальной полукруглой арки. Именно эти горизонтальные арки передают усилие распора массивной стене. Таким образом, каждый коноид как бы зажат между спандрелем, плоскостью стены и консольным выступом.

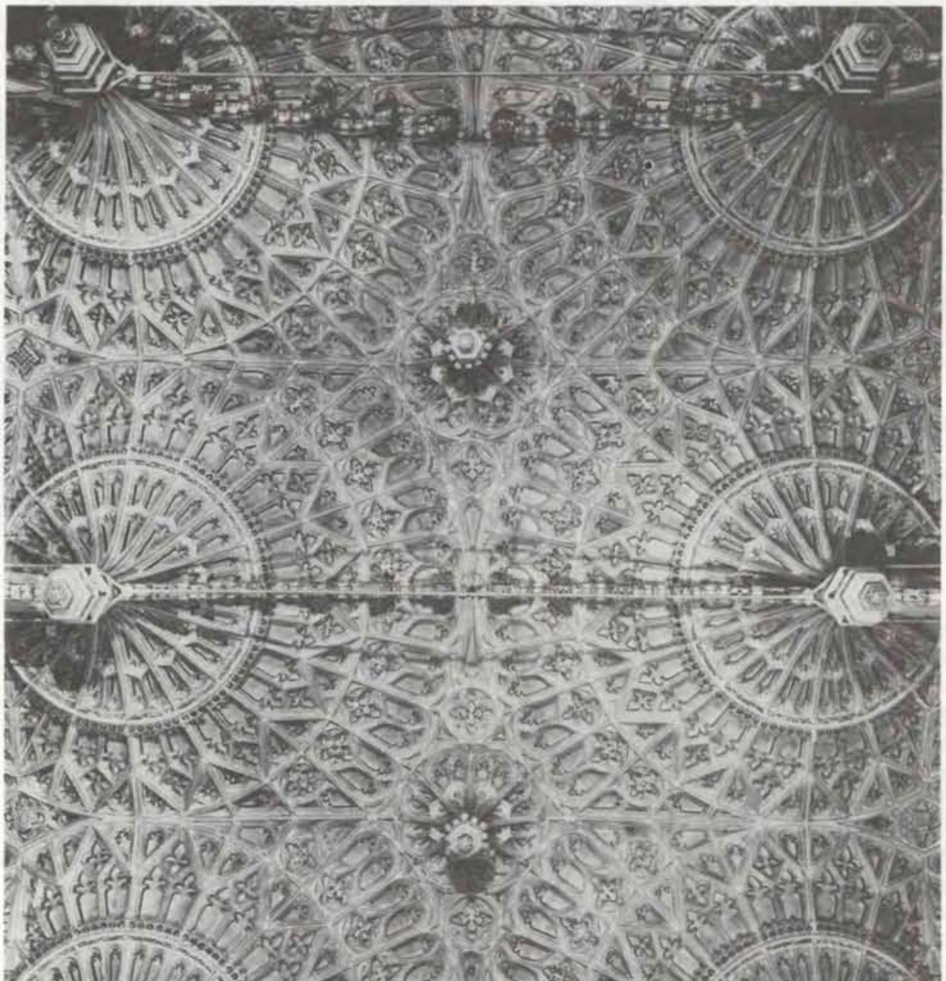
Усилия в коноиде не сосредоточены в ребрах, а относительно равномерно распределены; характер этого распре-

деления имеет принципиальное значение для статической устойчивости формы, поскольку коноид веерного свода представляет собой оболочку. Архитектурная оболочка определяется как напряженная структура, толщина которой мала по сравнению с ее поверхностью. Усилие, приложенное к произвольной точке оболочки, может быть представлено как равнодействующая сил, направленных в стороны и вниз; совокупность таких равнодействующих образует воображаемую поверхность, именуемую равнонапряженной. Математические расчеты показывают, что устойчивость коноида обеспечивается только в том случае, если эта воображаемая поверхность проходит в толще его физической оболочки, в противном случае возможно его разрушение.

Создатели веерных сводов не могли знать этого фундаментального закона. Анализ напряжений в оболочке выходил далеко за пределы их теоретических познаний, однако, опираясь на свой богатый опыт, они разработали ряд технических приемов, значительно повышавших устойчивость сводчатой конструкции. Узкая воронка внутри коноидальной оболочки над опорной консолью, именуемая карманом свода, в большинстве случаев заполнялась на высоту примерно одного метра обломками камня и заливалась известковым раствором. В капелле Кингз Колледж высота заполнения этой догрузочной массой почти одинакова для всех коноидов, поэтому можно сделать вывод, что этот технический прием был обдуман заранее.

Такого рода заполнение* обладает рядом достоинств. Во-первых, поверхность распределения усилий могла теперь выйти за пределы оболочки, в этом случае нагрузка приходилась уже не на узкую полосу примыкания оболочки к стене, а на значительно бóльшую ее поверхность. Во-вторых, заполнение карманов ведет себя как консольная нагрузка, уменьшающая распор в углах пролета, а тем самым и боковое давление на стены. Наконец, что наиболее важно для повышения устойчивости коноидов, забутовка уменьшает пролет, а тем самым и поверхность свода-оболочки. Нижняя, заполненная часть коноида является уже не тонкой оболочкой, а объемным телом, через которое проходит воображаемая равнонапряженная поверхность. Жак Хейман из Кембриджского университета показал, что оболочкой можно считать только ту часть веерного свода, которая располагается между обрезом спандреля и верхним уровнем забутовки. Эта часть свода несколько меньше, чем его общий пролет. Уменьшение поверхности оболочки приводит к увеличению отношения ее толщины к площади ее поверхности, поэтому вероятность, что равнонапряженная поверхность окажется в пределах между внешней и внутренней поверхностями оболочки, значительно возрастает.

* В русской терминологии — забутовка. *Прим. перев.*



КРУЖЕВНАЯ РЕЗЬБА потолков капеллы Генриха VII относится к числу лучших образцов декора веерных сводов. Капелла предназначалась для прославления династии Тюдоров, поэтому особая роль отводилась орнаменту. Своды сооружены полностью из заранее пригнанных каменных блоков, что гарантировало возможность выполнения сложнейшей декоративной резьбы.

ИТАК, не обладая достаточными теоретическими познаниями, средневековые мастера сумели обеспечить необходимую устойчивость веерного свода. При гарантированной устойчивости в принципе сохраняется риск обрушения конструкции из-за непрочности самого материала, но, так как естественный камень обладает высокой прочностью на сжатие, разрушение конструкции по этой причине маловероятно независимо от размера перекрытия. Нет принципиальных препятствий возведению большого свода по той же статической схеме, что и малого. И все же потребовался опыт многих десятилетий, чтобы строители убедились, что устойчивость веерного свода не связана с его размером. К концу XV в. этот факт становится известным, и английские зодчие приступают к реализации замыслов сооружения крупнейших веерных сводов.

Задача эта требовала осуществления весьма обширной строительной программы. Так, камень для перекрытия капеллы Кингз Колледж нужно было доставлять к строительной площадке по реке Кэм или на телегах. Добытые в карьере каменные глыбы имели размер порядка 70 см в поперечнике. В непосредственной близости от стройки располагались временные мастерские, в которых производилась обработка этих глыб.

Таким образом, своды капеллы сооружались из элементов, предваритель-

но обработанных в мастерских. Каменные блоки доводились в них до требуемого размера и формы, а поверхности будущих стыков между блоками здесь же тщательно выравнивались. Однако профиль ребер и резной декор на лицевой поверхности камней — элементов будущих коноидов намечались лишь грубо. Каждый камень получал маркировку, обозначающую квадрант пролета, в котором он будет находиться. У краев камня рядом с будущим швом вырезались арабские или римские цифры для обозначения горизонтального ряда кладки коноида, места данного камня в этом ряду и очередности его укладки.

Таким образом, каменщику, ведущему кладку, для точной установки камня нужно было только проверить его маркировку. Предварительная обработка камней свода требовала даже большего мастерства, чем их окончательная установка. Изучая маркировку, различимую на поверхности коноидов капеллы Кингз Колледж, я установил, что работы выполнялись в следующем порядке. Сначала возводились стены и устанавливались конструкции высокой кровли, стандартной для средневековой строительной практики, затем — поперечные арки, разделяющие отдельные пролеты зала, что требовало сооружения сложных временных кружал. Следующий шаг состоял в кладке коноидов, начиная от поперечных арок с последующим продвижением внутрь. Все коноиды сооружались одновременной кладкой их горизонтальных рядов. Когда кладка между арками смыкалась вдоль центральной оси пролета, на место устанавливался крупный замковый камень, связывающий две противолежащие арки. После завершения коноидов устанавливались камни продольного гребня, граничащие с боссом, а затем и сам босс.

Босс опускался на место сверху. За счет его большого веса в своде возникали дополнительные усилия сжатия, слегка изменявшие его форму, что позволяло легко разобрать верхний уровень лесов. Затем выполнялась окончательная резная обработка поверхностей. После заделки мелких случайных трещин коноиды нередко окрашивались. Своды капеллы Кингз Колледж предполагалось тщательно загрунтовать пигментом и покрыть позолотой.

САМЫЕ роскошные своды, несомненно, у капеллы Генриха VII, хотя по размеру своего пролета, равному 10,6 м, они несколько уступают сводам капеллы Кингз Колледж, имеющим наибольший пролет. Общее художественно-конструктивное решение перекрытий капеллы Генриха VII вызывает восхищение гармонией простого пространственного рисунка и сложного орнамента поверхности. Ее своды полностью собраны из заранее пригнанных элементов, иначе не удалось бы достичь такого великолепия резного декора, каковым должна была отличаться капелла монарха.

Конструктивное решение капеллы Генриха VII уникально. Мощные и большей частью скрытые в толще свода поперечные арки несут на себе расположенные ближе к стенам подвески. От этих подвесок «вырастают» коноиды свода. Поперечные арки, устремляясь вверх, как бы пронизывают коноиды от подвески, исчезая из виду в центральной части свода. Тем самым практически вся нагрузка от свода передается этим аркам на уровне крепления подвесок. Арки в свою очередь передают нагрузку стенам и усиливающим их контрфорсам. Таким образом, сводчатое перекрытие является структурным соединением оболочки (коноиды) и каркасной конструкции (поперечные арки).

В отличие от коноидов в большинстве веерных сводов, горизонтальные сечения которых представляют собой лишь дугу окружности, в капелле Генриха VII центральные коноиды имеют окружность в горизонтальных сечениях. Это вполне понятно, потому что круг — символ солнечного диска и обращения небесных сфер — играл важную роль в политической иконографии Тюдоров. В 1500 г., когда создавался проект капеллы, Тюдоры были молодой династией, всемерно стремившейся к самоутверждению. Роскошный резной декор перекрытий капеллы с множеством кольцевых форм должен был служить апофеозом земной славы монарха — избранника неба, каковым хотел предстать в истории Генрих VII. Потолки капеллы говорят нам более о короле и его дворе, чем о своем создателе, и мы можем только гадать о его личности.

Перекрытие капеллы Генриха VII — вершина развития веерного свода. В 1540 г. Генрих VIII положил конец власти и богатству монашеских орденов, в аббатствах которых возведено большинство сводов этого типа. Их строительство было приостановлено почти на столетие, пока его не возобновила англиканская церковь.

Сооружение веерных сводов не прекратилось полностью — в ряде новых соборов перекрытия имитируют славный древний стиль, однако этим запоздалым примерам недостает очарования тех сводов, что были созданы в период между 1350 и 1540 г. Они поражают не только своим великолепием, но и мастерством зодчих, которым в то время приходилось полагаться только на собственный опыт. Созданные ими своды отличаются не только совершенством форм — главное, к чему они стремились, — но и высокой прочностью.

Более точные значения числа π

ЧИСЛО π — отношение длины окружности к ее диаметру — приблизительно равно 3,14, а более точное может содержать любое количество дополнительных десятичных знаков. В конце XVI в. π уже умели вычислять до 30-го десятичного знака. С тех пор многие вычисляли более точные значения этого числа, продлевая его дробную часть. Почти на протяжении ста лет самым точным было значение π, содержащее 707 знаков после запятой. Его вручную рассчитал английский математик У. Шанкс, затратив на это 20 лет. Результат его работы был опубликован в 1873—1874 гг. в «Трудах Королевского научного общества». (В 1945 г. выяснилось, однако, что Шанкс допустил ошибку в 527-м знаке.) С появлением ЭВМ количество десятичных знаков числа π стало быстро увеличиваться — до 100 000 в 1962 г. и до 1 000 000 в 1973 г. В наше время два японских специалиста Йошиаки Тамура из Международной обсерватории в г. Мизусаве и Ясумаза Канада из Токийского университета рассчитали значение π с точностью до 2^{21} (2097152), 2^{22} (4194304) и 2^{23} (8388608) знаков после запятой.

На эту работу ушло несколько недель. Расчеты производились на быстродействующих цифровых ЭВМ по алгоритму, составленному несколько лет назад Ю. Саламином из Станфордского университета в Калифорнии. Тамура рассчитал π до более чем 2 000 000-го знака за 7,2 ч машинного времени. Затем, используя более быстродействующую машину Канада и Тамура вместе за 2,9 ч машинного времени провели расчеты, получив более 4 000 000 знаков после запятой. На еще более скоростной ЭВМ HITAC M-280H они за 6,8 ч получили 2^{23} десятичных знаков. Канада и Тамура намерены продолжать расчет до 2^{24} (16777216) знаков.

Главный побудительный мотив, который стимулирует проведение таких расчетов, объясняется стремлением найти закономерность в последовательности десятичных знаков в значении числа π. Широко распространено мнение, что последовательность знаков дробной части π никак не упорядочена, но это предположение не доказано. Если даже последовательность цифр в числе π действительно случайна, можно все же ожидать, что где-то существует, хотя и небольшая, локальная упорядоченность. Например, интересно выяснить, существует ли в числе π последовательный ряд 123456789. И если да, то где?

Термические осцилляторы

ДЖИРЛ УОЛКЕР

ТЕРМИЧЕСКИЙ осциллятор — интересное устройство, которое демонстрируют в школе на уроках физики как пример самоподдерживающейся нелинейной осциллирующей системы; одна из ее частей постоянно колеблется вследствие теплопередачи. Я расскажу о четырех таких устройствах, два из которых новые, а два других демонстрировались еще в XIX в. Хотя их действие основано на одном и том же принципе периодического перераспределения тепловой энергии, каждое устройство отличается от другого.

П. Виландер из Вашингтонского университета (г. Сиэтл, шт. Вашингтон) сконструировал термический водный осциллятор. Открытые концы U-образной трубки присоединяют к широким емкостям. Трубку полностью и каждую емкость примерно наполовину наполняют водой. Вверху емкости остаются открытыми. Нижнюю часть трубки нагревают, а верхние части охлаждают. Постепенно уровень воды в системе начинает колебаться между двумя ее частями с периодом от нескольких минут до нескольких часов.

Предположим сначала, что система находится в равновесии: обе ее части (правая и левая) содержат равные количества воды с одинаковым распределением температуры. Рассмотрим небольшой объем в тонком поперечном слое внизу трубки. Этот слой испытывает давление воды с обеих сторон.

Массу водного столба определяют два фактора — его высота и плотность воды. Плотность зависит от температуры: холодная вода имеет более высокую плотность, чем теплая. Поскольку сначала обе части системы идентичны как по уровню воды, так и по распределению температуры, рассматриваемый тонкий слой испытывает одинаковое давление с обеих сторон и остается в стационарном состоянии; при этом система находится в равновесии.

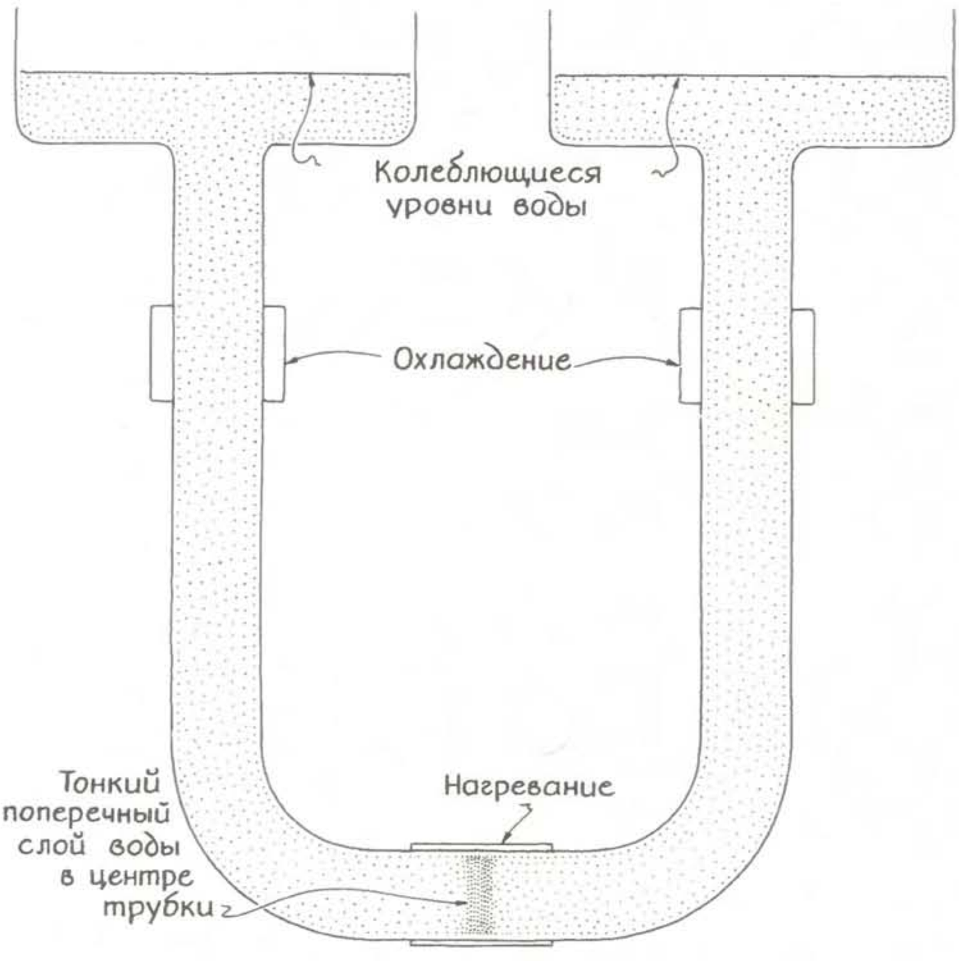
Предположим теперь, что внешнее возмущение нарушает равновесие, вызывая небольшое перетекание воды из левой части в правую. Симметрия системы нарушается: теперь ее части различаются как по уровню воды, так и по распределению температуры. Уровень воды в левом колене трубки снижается, и более холодная вода перетекает вниз. Соответственно уровень воды в правом колене повышается, и теплая вода в нижней части трубки перетекает вправо.

Если емкости вверху трубки достаточно широки, то сначала изменение уровней будет небольшим. В данном случае имеет значение нарушение симметрии распределения температуры. Поскольку плотность более холодной воды (в нижней части трубки слева) выше, чем более теплой (в нижней части трубки справа), давление на тонкий слой воды внизу трубки будет неодинаковым: давление с левой стороны становится больше. В результате уровень воды в правом колене увеличивается, а более теплая вода внизу трубки смещается вправо.

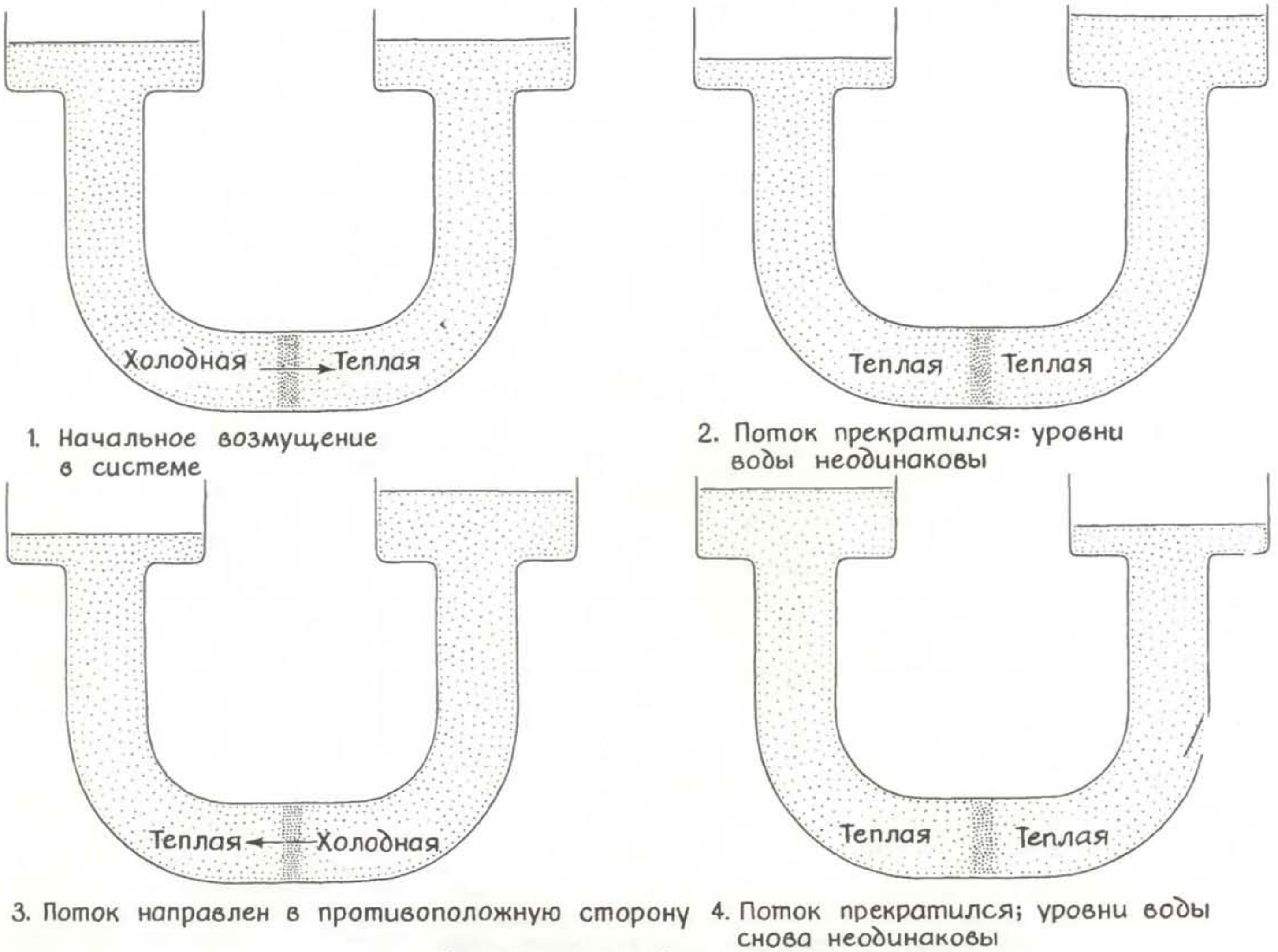
Хотя начальное изменение уровней мало, поток увеличивает эту разность до тех пор, пока она не станет довольно значительной. Тогда столб воды в правой части начинает оказывать большее давление на рассматриваемый слой. Поток воды в направлении слева направо замедляется, а затем прекращается, однако путем нагревания и охлаждения трубки можно восстановить симметрию распределения температуры. Когда поток прекращается, система приобретает примерно симметричное распределение температуры, но теперь в правой части уровень воды выше, чем в левой. Вода начинает перетекать в левую часть, цикл возобновляется в противоположном направлении, и система продолжает совершать колебания.

Осциллирующий поток существует неопределенно долго вследствие одновременного нагревания и охлаждения трубки. Если же вода достигнет одинаковой температуры, то колебания вскоре затухнут, и наступит стабильное состояние с равными уровнями воды в правой и левой емкостях. Однако колебания будут продолжаться до тех пор, пока происходит нагревание и охлаждение. Каждый раз, когда система проходит через равновесное состояние, при котором уровни воды одинаковы, разность плотностей обусловливает движение воды до тех пор, пока уровни станут разными.

Возникновение колебаний зависит от размера системы. Площадь поверхности воды в каждой емкости должна быть больше площади поперечного се-



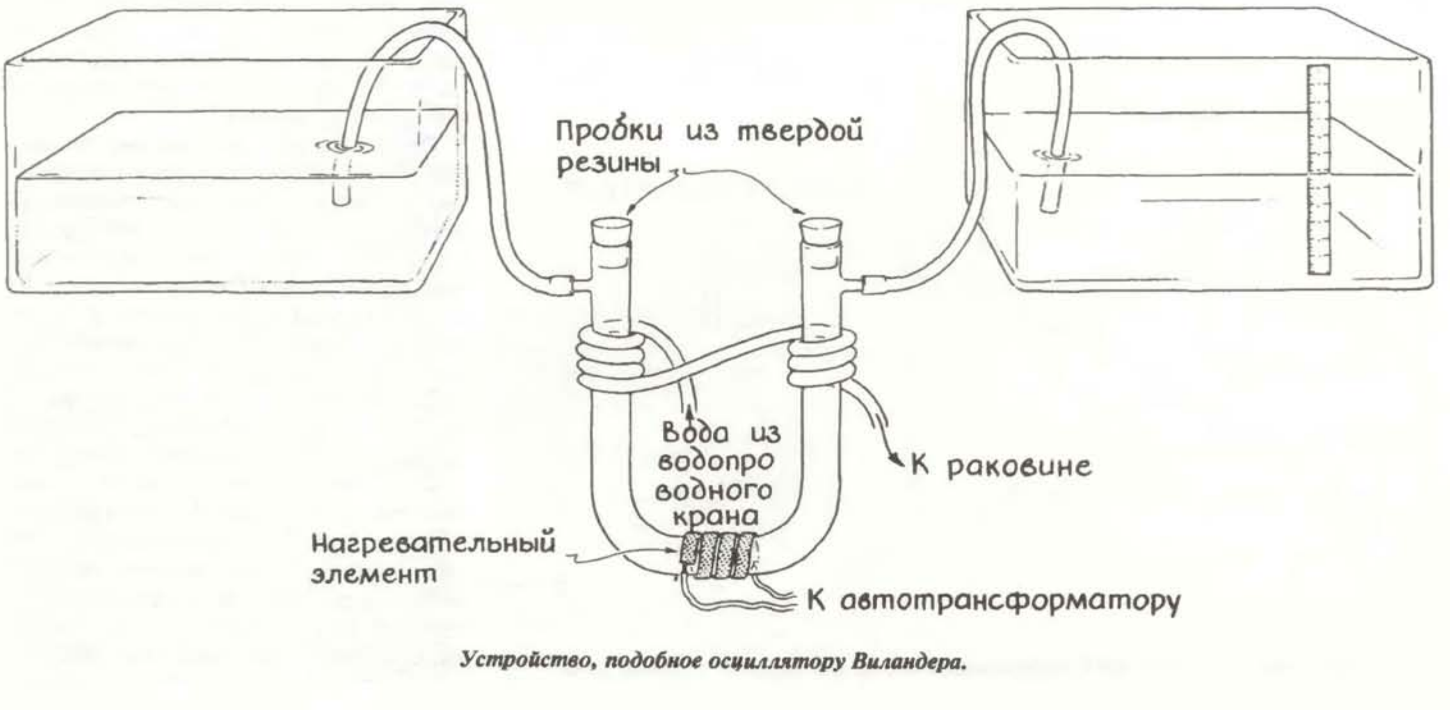
Физические условия для термического осциллятора, сконструированного П. Виландером.



Структура потока в водном осцилляторе.

чения трубки на величину, которая связана с распределением температуры. Отношение площади поверхности емкости к площади поперечного сечения трубки должно превышать значение, получающееся в результате деления числа 5000 на максимальную разность температур (в градусах Цельсия).

Например, если разность температур в системе составляет 50° С, то площадь поверхности воды в емкости должна быть по крайней мере в 100 раз больше площади поперечного сечения трубки. Для меньшей разности температур это отношение должно быть еще больше. Если емкости очень большие, то колебания уровня воды трудно заметить.



Устройство, подобное осциллятору Виландера.

Требование об относительных размерах указанных площадей обусловлено влиянием давления столба жидкости в центральном объеме внизу U-образной трубки. Предположим, что вследствие начального внешнего возмущения небольшое количество воды перетекает вправо. Возникающий перепад температур способствует перемещению в том же направлении еще большего количества воды. Этому движению противодействует разность уровней воды в обеих частях системы. Если площадь поверхности емкостей очень мала, то перетекание сразу же создает такую разность уровней воды, которой более чем достаточно для того, чтобы остановить поток, возникающий вследствие асимметрии распределения температуры. В этом случае равновесие в системе быстро восстанавливается.

Чтобы началось колебание, площадь поверхности воды в емкостях должна быть относительно большой. В этом случае поток, вызываемый асимметрией распределения температуры, не создает сразу большой разности уровней воды; он развивается постепенно. Период колебания определяется такими факторами, как отношение площадей поверхности воды в емкостях и поперечного сечения трубки, длина трубки, показатель вязкости воды и сила тяжести.

Аппарат, подобный тому, который сделал П. Виландер, можно изготовить, припаяв стеклянные емкости к концам U-образной трубки. Обычно берут трубку из жаропрочного стекла для того, чтобы она не лопнула при нагревании и охлаждении. Но для меня припаять стеклянные емкости к такой трубке представлялось слишком сложной задачей. К тому же я не знал, каким образом можно установить и закрепить всю систему в вертикальном положении. Поэтому я выбрал U-образную трубку с боковыми отростками около верхних краев и соединил их коротким резиновым шлангом с пластмассовыми стаканами. Стаканы имели одинаковую емкость, и площадь поверхности каждого из них была в 200 раз больше площади поперечного сечения U-образной трубки.

Я наклеил полоску миллиметровки на внутреннюю поверхность одного из стаканов так, чтобы можно было измерять уровень воды. Стеклянная трубка полностью, а каждый стакан примерно наполовину были наполнены водой. Вокруг нижней части трубки размещался обычный спиральный электронагреватель, который подключался а автотрансформатору для того, чтобы можно было регулировать степень нагревания. (Необходимо проследить, чтобы не было протечки воды и ее контакта с электропроводом.)

Для охлаждения стеклянной трубки я использовал тонкостенный резиновый шланг, соединенный с водопроводным краном. Хотя столь примитивным способом действительно удавалось охладить воду в верхней части трубки, мне кажется, что нужна была более совершенная конструкция. Теплопередача через стеклянные стенки трубки и резиновый шланг, по-видимому, происходит очень медленно.

Я не знал, каким образом измерить распределение температуры воды в трубке. При разных положениях ручки автотрансформатора мне удалось заметить слабые колебания уровня воды в стаканах. Уровень поднимался или опускался примерно на миллиметр каждые пять минут. Колебания были довольно слабые и неравномерные; когда прекращалось нагревание и охлаждение системы, они также прекращались.

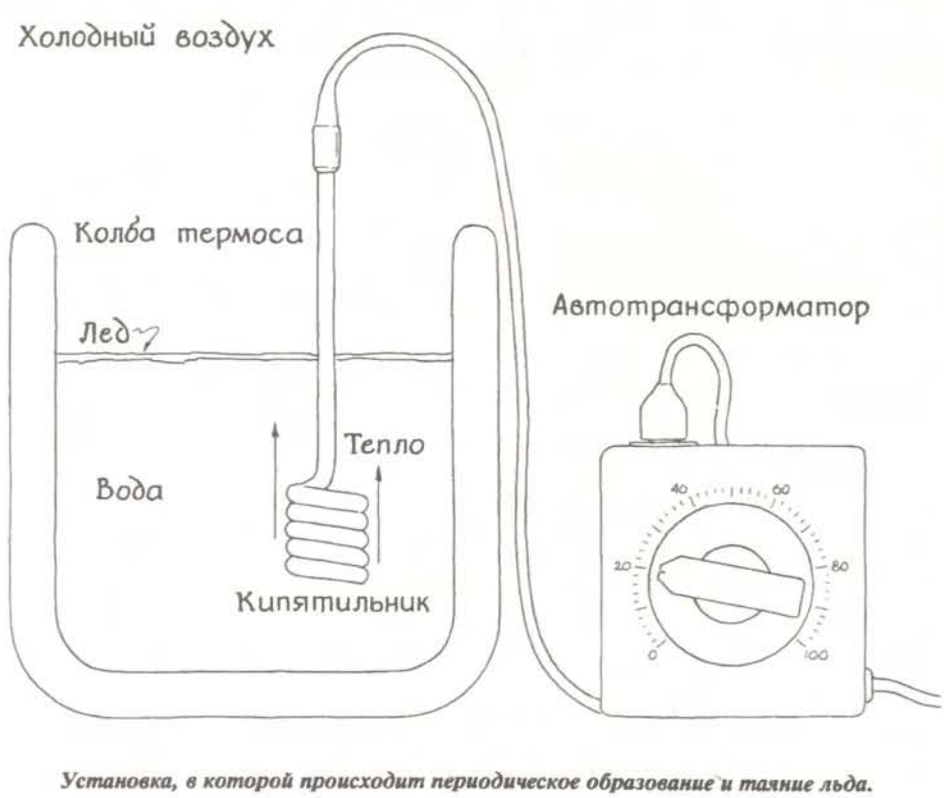
П. Виландер посоветовал мне, как сделать более чувствительный индикатор уровня воды. Маленькую пробку свободно подвешивают на короткой проволочке за край стакана; пробка должна плавать на поверхности воды. Сверху в пробку вставляют соломинку в вертикальном положении. Когда уровень воды изменяется, плавающая пробка начинает вращаться вокруг своей оси, а верхний конец соломинки увеличивает момент движения.

П. Виландер описал также еще один тип термического осциллятора, в котором периодически образуется тонкий слой льда. Он использовал цилиндрический сосуд из органического стекла, наполненный 50 л воды; длина окружности его дна составляла 50 см. Сосуд был теплоизолирован; к его дну был присоединен нагревательный элемент мощностью примерно 100 Вт. Магнитные мешалки постоянно перемешивали воду для того, чтобы обеспечить равномерное распределение тепла. Вся конструкция помещалась в морозильную камеру, где температура воздуха поддерживалась на уровне —20 °С. Воздух над поверхностью воды также постоянно перемешивался.

П. Виландер регистрировал изменение температуры воды на половине ее глубины. Периодическое образование льда на поверхности воды и его таяние происходили в интервале температур примерно 10 °С. Когда поверхность воды замерзала, ее температура повышалась примерно до 10 °С. При таянии льда температура в системе падала почти до 0 °С. Период колебания составлял примерно 50 ч.

Периодическое образование и таяние льда можно наблюдать и в домашних условиях. П. Виландер предлагает использовать для этого колбу термоса с широким горлом. Наполните ее водой почти доверху, опустите нагревательную спираль и поместите в холодильную камеру. Лед будет периодически образовываться и таять, хотя колебания не будут такими равномерными, как те, которые наблюдал П. Виландер в аппарате большего размера.

Я следовал его рекомендациям, но в качестве нагревательного элемента использовал обычный кипятильник. Кипятильник был подвешен в центре термоса и закреплен с помощью прочной клейкой ленты. Однако верхняя часть кипятильника выступала над поверх-



Установка, в которой происходит периодическое образование и таяние льда.

ностью воды, и, конечно, выделяемое им тепло мешало периодическому образованию льда.

Я присоединил кипятильник к автотрансформатору и мог регулировать скорость нагревания воды. Чтобы можно было плотно закрыть дверь камеры, прижав шнур от автотрансформатора, резиновая прокладка на двери должна быть достаточно эластичной. Регулятор морозильной камеры я поставил на отметку максимального холода. Как и в предыдущем случае, все было предусмотрено для того, чтобы вода случайно не попала на электропровод.

Периодическое образование льда объясняется худшей теплопередачей через лед по сравнению с теплопередачей через воду, которая постоянно перемешивается либо механически, либо в результате конвекции. Тепло, поступающее в систему от кипятильника, должно подниматься вверх, потому что стенки колбы термоса теплоизолированы. Поверхность воды охлаждается воздухом в морозильной камере, температура которого гораздо ниже точки замерзания воды.

Система может находиться в двух возможных стабильных состояниях. В одном из них вода на поверхности остается жидкой, поскольку воздух не особенно холодный, а тепло, поступающее от кипятильника, постоянно мешает замерзанию. Вследствие теплопроводности и конвекции тепло передается вверх, причем достаточно быстро, так что лед не успевает образоваться.

В другом стационарном состоянии поверхность воды постоянно замерзшая, поскольку воздух очень холодный, а скорость подвода тепла не слишком высока. Тепло от нагревательного элемента передается вверх к слою льда за счет теплопроводности и конвекции. При этом оно медленно проводится через лед и уносится конвекционными потоками воздуха. Система будет стабильной, потому что, хотя теплопроводность льда низкая, она достаточна для того, чтобы температура не могла подняться до точки таяния льда.

Колебания могут начаться, если воздух достаточно холодный, а скорость подвода тепла от нагревательного элемента высока. Предположим, что вода не замерзла. Холодный воздух забирает тепло с поверхности воды достаточно быстро, для того чтобы началось замерзание. Когда образуется слой льда, отвод тепла из системы замедляется, потому что лед не так хорошо проводит тепло, как перемешиваемая вода. Температура воды подо льдом начинает повышаться, и постепенно вода нагревается настолько, что лед тает.

Когда лед исчезает, скорость теплопередачи увеличивается и температура воды падает. Затем температура поверхности снова достигает точки замерзания и образуется новый слой льда.

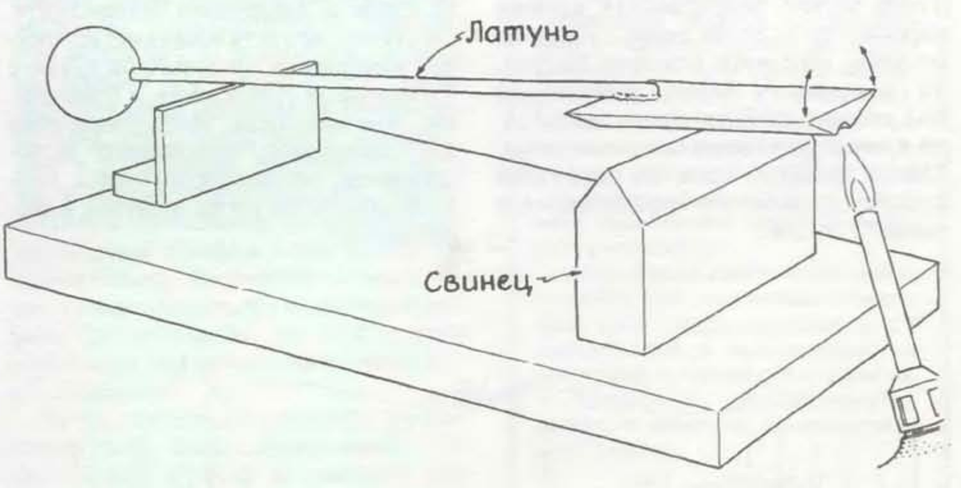
Термический осциллятор другого типа был открыт в 1805 г. инспектором плавильного завода в Саксонии. Когда нагретый металлический стержень кладут поперек бруска из другого металла, стержень может сильно вибрировать, издавая шум. Независимо осциллятор такого же типа примерно 25 лет спустя сконструировал А. Тревельян, и с тех пор это устройство известно как «коромысло Тревельяна», или «качалка Тревельяна».

Я видел два варианта осциллятора этого типа. В каждом случае очень важен выбор металлов. В первом случае две свинцовые пластины зажимают в тиски. Верхние края пластин, выступающие из тисков, остро заточены так, что между ними образуется небольшой зазор. Поперек пластин помещают латунный стержень таким образом, что он находится в равновесии. Когда его нагревают, он — как коромысло, лежащее на точке опоры — начинает качаться, ударяясь то об одну, то об другую пластину; при этом раздается почти ритмичное постукивание.

Во втором случае берут латунный брусок треугольного сечения с треугольной канавкой вдоль одного ребра, превращающей ребро в два параллельных. Оба этих ребра лежат на кромке свинцового блока. От латунного бруска треугольного сечения отходит на одном уровне с двойным ребром стержень, несущий груз на дальнем конце; этот конец стрежня свободно лежит на опоре. При нагревании бруска его острозаточенные кромки покачиваются, ударяясь о свинцовый блок, и издают звук приблизительно постоянной частоты.

В устройстве Тревельяна стержень был медным. Сейчас обычно используют латунь. Вероятно, вместо латуни, меди или свинца можно взять другие металлы. Однако, чтобы добиться постоянных колебаний, металлы нужно подбирать очень тщательно. Принцип действия такого осциллятора впервые объяснил Дж. Лесли вскоре после того, как Тревельян продемонстрировал свою «качалку». Более поздние работы лишь уточнили это объяснение.

Рассмотрим устройство, в котором латунный брусок свободно лежит на свинцовом блоке. Районами контакта являются два острых ребра в канавке бруска. Когда латунный стержень и латунный брусок нагревают горелкой, тепло передается к свинцу через эти два ребра. Скорость теплопередачи за-





Два варианта качалки Тревельяна.

висит от веса, приходящегося на одно ребро: бо́льший вес обеспечивает лучший контакт и лучшую теплопередачу. Одно из ребер сначала передает тепло лучше. Тепловое расширение свинца вызывает появление небольшого бугорка в том месте, к которому прикасается ребро. Распространение тепла может быть достаточно быстрым, для того чтобы сместить ребро вверх, в сторону от возникшего бугорка.

Здесь происходят два явления. Поскольку одно ребро чуть-чуть приподнято, вес латунного бруска переносится на другое ребро. Тепло тоже передается к свинцу через второе ребро. Бугорок под первым ребром охлаждается и оседает. А под вторым ребром возникает новый бугорок. Система совершает колебания, в которых повторяются циклы образования и исчезновения бугорков за счет расширения свинца.

Частота постукивания в какой-то степени зависит от размеров коромысла. В некоторых случаях она настолько высока, что появляется непрерывный звук.

Почему не все материалы дают такой эффект? Первое условие успешного опыта заключается в том, что блок должен быть из материала, который расширяется при нагревании. Свинец расширяется хорошо, даже в большей степени, чем латунь и медь. Второе условие состоит в том, чтобы тепло, передаваемое к блоку через ребро, распространялось не слишком быстро. Тогда можно поддерживать качания коромысла. Если же тепло отводится от точек контактов слишком быстро, то поверхность блока, находящаяся под ребром, не будет иметь достаточно времени для теплового расширения. Свинец проводит тепло как раз в такой степени, в какой это необходимо для данного опыта.

И наконец, третье условие — блок должен проводить тепло достаточно быстро, чтобы тепловое расширение не продолжалось слишком долго, а бугорок не слишком медленно исчезал. В противном случае ребро не будет приподниматься над блоком.

Латунь и свинец можно поменять местами таким образом, чтобы уже свинцовый брусок с двумя острыми ребрами покоился на латунном блоке, но по-прежнему нагревать надо латунь. В такой системе колебания легче наблюдать, если освещать коромысло стробоскопической лампой. Слабые колебания можно усилить, приделав иглу или тонкую проволочку к коромыслу под прямым углом к его длинной оси. Конец иглы будет проходить большее расстояние, чем само коромысло, что облегчает наблюдение колебаний.

Еще один термический осциллятор демонстрировался когда-то на уроках физики в школе. Кусок проволочной сетки монтируют примерно на четверти длины от нижнего конца внутри вертикальной трубы. Сетка загибается по краям так, чтобы она удерживалась в определенном месте внутри трубы. После этого под трубу подводят горелку для нагрева и трубы, и сетки в ней. Когда горелку убирают, труба гудит в течение нескольких секунд. Звуковой тон получается почти чистым вследствие определенной частоты звуковых колебаний в трубе.

Эта система была открыта в 1859 г. П. Рийке в Лейденском университете. Он считал, что звук возникает вследствие сокращения проволочной сетки и трубы при их охлаждении, и предполагал, что колебания, сопровождающие эти сокращения, возбуждают звуковую волну, исходящую от трубы. Рийке ошибался: ни сетка, ни труба не колеблются.

Точное объяснение этого явления было дано Рэлеем в 1878 г. Звук появляется благодаря возникающим в трубе колебаниям воздуха. Когда горелка нагревает воздух в трубе, конвекционный поток направлен вверх, через сетку вдоль трубы. При этом не возникает звуковых колебаний, поскольку теплый воздух движется в одном направлении — вверх.

Когда же источник тепла убирают, конвекционный поток некоторое время продолжается, поскольку и труба, и сетка нагреты. Возмущения в столбе воздуха в трубе создают звуковую стоячую волну. Сначала я объясню, что такое стоячая волна в ненагретой трубе, а затем покажу, как конвекционный поток в трубе Рийке усиливает эту волну и труба начинает «гудеть».

Звуковая стоячая волна внутри прямой трубы может возникнуть при взаимодействии двух звуковых волн, распространяющихся по трубе в противоположных направлениях навстречу друг другу. Эти звуковые волны должны быть одинаковыми, за исключением направления движения. Они проходят по трубе, отражаются от ее открытых концов, а затем снова проходят по трубе. Когда они налагаются друг на друга, их интерференция вызывает слабые колебания частиц воздуха, направленные вдоль трубы. В результате интерференции возникает устойчивая картина колебаний частиц воздуха. В некоторых сечениях вдоль трубы частицы вообще не двигаются: такие точки называют узлами волны, в других же колебания максимальны: эти точки называют пучностями.

Описанную картину колебаний называют стоячей волной, потому что положения ее узлов и пучностей неподвижны. Простейшую стоячую волну в трубе называют основной: ее узел находится в середине трубы, а пучности — у обоих ее концов. Таким образом, около открытых концов колебания частиц воздуха в направлении вдоль трубы максимальны. В то же время около центра трубы эти колебания очень слабые или вообще частицы воздуха не движутся.

Чтобы возбудить основную волну в трубе, распространяющиеся звуковые волны должны иметь частоту, равную скорости звука, деленной на удвоенную длину трубы. Как только стоячая волна инициирована, ее амплитуда может возрастать, если увеличивается энергия звуковых волн. В этом случае говорят, что труба резонирует с частотой основной стоячей волны. Когда достигается резонанс, звук из трубы может быть довольно громким.

Звуковая стоячая волна также вызывает периодические изменения давления воздуха вдоль продольной оси трубы. Представим себе, что узел находится в центре трубы, когда основная



Движения воздуха в трубе осциллятора Рийке.

волна уже образовалась. С каждой стороны от узла частицы воздуха колеблются в направлении вдоль трубы. В первой фазе колебания частицы, находящиеся на противоположных сторонах от узла, движутся по направлению к его центру, повышая давление воздуха в этой точке, а во второй фазе — в направлении от узла, создавая там зону низкого давления. В процессе колебаний давление в узле становится то выше, то ниже атмосферного давления вне трубы. Таким образом, самые большие изменения давления внутри трубы происходят в узле волны.

Узлам и пучностям звуковой волны соответствуют минимумы и максимумы звукового давления. Объяснение, данное Рэлеем для опыта с трубой Рийке, основано на рассмотрении относительного положения проволочной сетки и узла волны (минимума звукового давления) в центре трубы. Труба должна находиться в вертикальном положении, чтобы тепло вызывало в ней конвекционный поток. Сетка должна быть расположена в нижней части трубы, не слишком близко к нижнему ее концу или к узлу волны в центре трубы.

Предположим, что разные возмущения массы воздуха внутри трубы при ее охлаждении инициируют слабые волны, которые порождают основную стоячую волну. В первой фазе колебаний частицы воздуха движутся вверх через сетку по направлению к узлу, увеличивая давление в узле, а во второй — вниз через сетку в направлении от узла, создавая в узле пониженное давление.

Рэлей доказал, что колебаниям воздуха внутри трубы способствует теплопередача от нагретой сетки и нагретой трубы. Это действительно эффективно, но только в одной — первой фазе цикла колебаний. Когда частицы воздуха движутся вверх через сетку к узлу волны, относительно холодный воздух засасывается в трубу через ее нижний открытый конец. Таким образом, отдаваемое сеткой тепло способствует движению частиц воздуха вверх по направлению к узлу, повышая давление в этой точке и интенсифицируя процесс колебаний.

Во второй фазе колебаний, при движении частиц воздуха вниз, тепло от сетки по-прежнему стремится направить их вверх. В этом случае частицы воздуха как бы подталкиваются в противоположном направлении, препятствующем их колебанию; однако сила такого воздействия невелика. Воздух, перемещающийся вниз, почти такой же горячий, как и сетка, так как он только что прошел через нее. Поэтому теплопередача от сетки в воздух здесь мала. То, что теплопередача действительно происходит, лишь ослабляет колебания, поскольку нагревание воздуха заставляет его двигаться только вверх, а не вниз (синхронно со звуковой волной).

Колебания воздуха в основной стоячей волне в трубе усиливаются асимметрией теплопередачи в этих двух фазах колебаний. Фактически колебания в одной фазе получают дополнительный импульс, а в другой нет. Рэлей проводил аналогию с силой, периодически прикладываемой к раскачивающемуся маятнику. Представьте, что вы подталкиваете маятник каждый раз, когда он проходит через самую нижнюю точку. В одной фазе качания подталкивание добавляет энергию маятнику, потому что вы подталкиваете его в направлении движения. Но при следующем проходе маятника через самую нижнюю точку вы будете толкать его уже в направлении, противоположном движению. Таким образом, вы как бы отнимаете у него энергию. Если прикладывать ту же силу каждый раз, когда маятник приходит в нижнюю точку, то в результате он не получает дополнительной энергии. Раскачивание постепенно замедляется вследствие фрикционных потерь энергии. Чтобы усилить раскачивание, нужно подталкивать маятник только тогда, когда он приобретает дополнительную энергию.

При теплопередаче от сетки к воздуху, в случае когда он перемещается к центру трубы, имеется еще одно преимущество. В середине трубы повышение температуры и давления воздуха происходит пропорционально. Когда нагретая сетка повышает температуру воздуха, она одновременно способствует и повышению давления в центре трубы. Если воздух начинает перемещаться в направлении от средней точки, понижая давление в ней, любое дополнительное нагревание воздуха будет снова стремиться увеличить давление. Следовательно, на этой стадии небольшая теплопередача способствует колебанию.

Чтобы продемонстрировать осциллятор Рийке, можно использовать любую трубу, которая не плавится при нагревании. А в качестве сетки можно взять даже кусок металлической оконной сетки. Двойной слой дает еще лучшие результаты, чем одинарная сетка. Нагревать трубу можно лабораторной горелкой, на газовой плите и даже свечкой.

Как только труба «загудит», попробуйте повернуть ее в горизонтальное положение, а потом снова верните в вертикальное. Когда труба находится в горизонтальном положении, естественная конвекция вдоль трубы нарушается и гудение сразу же прекращается. Если трубе придать опять вертикальное положение, пока сетка еще не остыла, гудение возобновляется.

Занимательный Лисп — язык искусственного интеллекта

ДУГЛАС Р. ХОФШТАДТЕР

МНЕ УЖЕ часто приходилось писать в своих статьях об искусственном интеллекте — области, в которой идет поиск методов программирования, позволяющих компьютерам демонстрировать такие качества, как гибкость, здравый смысл, элементы понимания и творчества, "самосознание" и чувство юмора. Искусственным интеллектом (ИИ) стали заниматься немногим более двадцати лет назад, и с тех пор уже возникли многочисленные направления, так что теперь — это область активных и разносторонних исследований. В США, вероятно, насчитывается около двух тысяч человек, профессионально занимающихся разработками, связанными с искусственным интеллектом, примерно столько же специалистов работают в этой области и в других странах. И хотя имеются значительные разногласия по поводу того, какой подход к ИИ следует считать наилучшим, в одном почти все единодушны — в выборе языка программирования. Большинство разработок ИИ было реализовано на языке, получившем название Лисп. Это название произошло от английских слов List Processing (Lisp) — обработка списков.

Почему большая часть исследований в области ИИ проводится на базе Лиспа? Для этого есть много причин в основном технического характера, однако одна из главных причин весьма проста и заключается в том, что Лисп имеет очень четко очерченную основу. Каждый язык программирования обладает какими-то произвольными свойствами, или конструкциями, и большинство языков в действительности даже перегружено ими. Лишь немногие языки, такие, как Лисп и Алгол, построены на основе некоторого ядра, кажущегося настолько же стройным и естественным, как одна из областей чистой математики. Ядро Лиспа отличается кристальной чистотой, которая не только делает этот язык привлекательным с эстетической точки зрения, но и придает ему гораздо большую гибкость по сравнению с большинством других языков. Лисп — красивый язык, и он занимает центральное место в искусственном интеллекте — одной из важных областей современной науки. Поэтому я решил посвятить ему серию из нескольких статей, где будут представлены некоторые основные конструкции этого языка.

Вообще говоря, своими корнями Лисп уходит в математическую логику. Такие пионеры математики, как Торальф Сколем, Курт Гёдель и Алонсо Чёрч в 20-е и 30-е годы нашего столетия привнесли в логику новые плодотворные идеи, которые десятилетиями позже воплотились в Лиспе. Писать программы для ЭВМ начали еще в 40-е годы, однако "языки высокого уровня" (к которым относится и Лисп) возникли только в 50-е годы. Первым языком для обработки списков был не Лисп, а IPL (сокращение от английских слов Information Processing Language — язык для обработки информации), разработанный в середине 50-х годов Г. Саймоном, А. Ньюэллом и Дж. Шоу.

В 1956-1958 гг. Дж. Маккарти, используя эти более ранние разработки, создал новый изящный алгебраический язык для работы со списками, которому он дал имя Лисп. Этот язык стал популярным среди молодых сотрудников Маккарти, участвовавших в разработке новой в то время темы искусственного интеллекта в Массачусетском технологическом институте (МТИ), он был реализован на машине IBM 704, его приняли в других исследовательских лабораториях по искусственному интеллекту, и все эти годы он широко использовался. В настоящее время существует уже много диалектов этого языка, однако все они построены вокруг одного и того же изящного ядра.

ПОСМОТРИМ теперь, как в действительности работает Лисп. Одна из наиболее привлекательных особенностей Лиспа заключается в том, что он интерактивный, т.е. работает в режиме диалога, что отличает его от большинства других языков высокого уровня, не обладающих этим свойством. На практике это выглядит следующим образом. Пожелав воспользоваться Лиспом, вы садитесь за терминал, соединенный с ЭВМ, в памяти которой находится система Лиспа, и набираете на клавиатуре терминала слово lisp (или какое-то другое стартовое слово — это зависит от соглашений, принятых для конкретной системы).

В ответ на экране терминала появится "сигнал ожидания", характерный символ; чаще всего в роли этого символа выступают стрелочка или звездочка. Мне, например, нравится представлять себе этот сигнал как приветствие, которым встречает вас волшебный джин Лиспа. Низко кланяясь, он как бы говорит вам: "Я здесь, чтобы выполнять ваши желания. Что вам угодно?". Затем джин ждет, пока вы не напечатаете ему какую-нибудь команду. На самом деле эту систему называют, конечно, не джином, а *интерпретатором*. Он выполняет любую вашу команду, однако, формулируя свое желание, вы должны соблюдать осторожность. Даже небольшая неточность может привести к роковым ошибкам.

Ниже показан сигнал ожидания системы Franz (один из диалектов Лиспа), он сообщает о том, что джин системы Franz готов выполнить ваше желание.

```
→
```

Итак, джин ожидает команды. Напечатаем ему простое выражение:

```
→ (plus 2 2)*
```

и нажмем на клавишу возврата каретки. Даже те, кто никогда не работал с Лиспом, наверное, догадаются, что в ответ джин выведет на экран символ "4". Затем он выдаст еще один сигнал ожидания, так что на экране появятся три строчки:

```
→ (plus 2 2)
4
→
```

Теперь джин готов выполнить нашу следующую команду, или, выражаясь вежливее, наше следующее желание. Выполнение желания, сформулированного в виде оператора Лиспа, называется *вычислением* этого оператора. Приведенный выше короткий диалог между человеком и компьютером дает нам представление о действиях интерпретатора Лиспа: он *читает* оператор, *вычисляет* его, *печатает* полученное значение и сигнализирует о своей готовности прочесть новый оператор. Поэтому о последовательности действий интерпретатора говорят как о цикле "чтение — вычисление — печать".

Лисп является интерактивным языком благодаря существованию джина Лиспа (интерпретатора). Как только вы напечатали свое желание — в форме оператора языка, — немедленно устанавливается "обратная связь" с

* В качестве "служебных слов" Лиспа, так же как и ряда других языков — Фортран, ПЛ/I, Алгол и т.п., — приняты английские слова, поэтому во фрагментах программ мы не будем их переводить, ограничиваясь лишь пояснением их значения в тексте. — *Прим. перев.*

Лиспом. И чтобы все ваши желания были реализованы, нужно сначала напечатать одно желание, попросить джина выполнить его, напечатать другое, снова обратиться к джину и т.д.

С другой стороны, при работе со многими другими языками высокого уровня (но не со всеми) вы должны написать целую программу, состоящую из большого количества "желаний", которые выполняются в строго установленном порядке. Еще хуже то, что последующие желания обычно очень сильно зависят от результата выполнения предыдущих, при этом вы, конечно, не можете следить за выполнением каждого отдельного желания — команды. Неудивительно поэтому, что выполнение подобной программы может привести к неожиданным результатам — слишком велико количество команд, которые должны действовать согласованно. Если вы допустили малейшую неточность при формулировании последовательности ваших желаний, то весьма вероятно, что ваша программа вообще не будет работать. Прогнать такую программу — это все равно что запустить совершенно неиспытанный космический корабль: вы просто не в состоянии предвидеть все ситуации, в которых что-то сработает не так, как надо, остается только ждать в надежде на то, что все будет в порядке. Если происходит срыв, вы ищете причину, вносите исправление и повторяете запуск. Этот неуклюжий, окольный, дорогостоящий способ программирования резко отличается от прямого, интерактивного стиля, присущего программированию на Лиспе, когда ваши желания выполняются одно за другим, что позволяет составлять и отлаживать программу в "аддитивной" манере. Это вторая важная причина популярности Лиспа.

КАКОГО ТИПА желания вы можете напечатать джину Лиспа, что он может вычислить и напечатать в ответ? Начнем с того, что ему можно задавать арифметические выражения, правда, записанные в несколько непривычной форме, например (times (plus 6 3) (difference 6 3)). Здесь times означает "умножить", plus — "сложить", а difference — "вычесть". Ответом будет 27, поскольку вычисление (plus 6 3) дает 9, (difference 6 3) дает 3, а произведение этих чисел равно 27. Такая запись арифметических выражений, или нотация, когда символ операции находится слева от операндов, была изобретена польским специалистом в области логики Яном Лукашевичем, когда ЭВМ еще не было и в помине. К сожалению, фамилия Лукашевич казалась слишком сложной в англоязычных странах и эту нотацию стали называть "польской". Теперь зададим вам простую задачку, записанную в польской нотации. Вы будете играть роль джина Лиспа:

```
→ (quotient
     (plus 21 13)
     (difference
         23
         (times
             2
             (difference 7 (plus 2 2)))))
```

Здесь quotient означает "разделить".

Вероятно, вы уже заметили, что в операторах Лиспа используются скобки. Изобилие скобок — одна из отличительных черт языка. Нередко встречаются выражения, заканчивающиеся дюжиной скобок, однако, когда привыкаешь к своеобразному виду выражений Лиспа, они, напротив, кажутся очень удобными для чтения и даже радуют глаз, особенно если записаны так, что выявляется логическая структура выражения. Эффект достигается за счет тщательно подобранных отступов и переходов на следующие строчки (как в выражении, приведенном выше).

Пожалуй, главное в Лиспе — это структуры, которыми он манипулирует. Работа любой программы заключается в создании, изменении и уничтожении структур. Они подразделяются на два вида — простые и составные, или, как их обычно называют, *атомы* и *списки*. Таким образом, любой объект Лиспа — это либо атом, либо список (но не то и другое одновременно). Единственное исключение составляет специальный объект nil, являющийся и атомом и списком. О нем мы поговорим ниже. Как выглядят типичные атомы Лиспа? Приведем несколько примеров:

водород, гелий, й-с-бах, 1729, пи, фу, бар, баз, кнопки-&-скрепки

СПИСКИ — ЭТО гибкие структуры данных Лиспа. По существу эта структура соответствует своему названию: набор некоторых элементов, расположенных в указанном порядке. Иногда эти *элементы* списка называют *членами* списка. Что они собой представляют? Во-первых, и это неудивительно, элементами списка могут быть атомы. Однако с таким же успехом в роли элементов списка могут выступать другие списки, элементами которых также могут оказаться списки и т.д., — здесь имеет место рекурсия. Стоп! Тут я уже забегаю вперед. Но ничего. Вы наверняка поняли, что я имел в виду, и теперь подготовлены к более точному определению этого слова, которое будет дано позже.

Список, напечатанный на экране дисплея, легко узнать по его скобкам. Все, что заключено в скобки (число открывающих скобок равно числу закрывающих), — это список. Например, (дзынь бах трах (бряк шмяк)) — список из четырех элементов, причем четвертый элемент сам представляет собой двухэлементный список. Еще один пример — короткий список (plus 2 2). Он показывает, что операторы Лиспа также являются списками. Последнее очень важно, поскольку это означает, что джин Лиспа, манипулируя списками и атомами, может сам составлять новые желания — команды. Поэтому ваше желание может, в частности, заключаться в том, чтобы построить и затем вычислить (т.е. реализовать) совершенно новое желание.

Существует также пустой список, не содержащий ни одного элемента. Как он записывается? Возможно, вам придет на ум форма "( )". И действительно, такая запись правильна, но есть и другой способ обозначения пустого списка, а именно: записать слово nil. Эти обозначения — синонимы, однако nil применяется чаще, чем "( )". Пустой список — одно из ключевых понятий в Лиспе, среди множества возможных списков он играет ту же роль, что нуль среди возможных чисел. Еще одно образное сравнение: nil подобен земле, в которой коренятся все структуры. Однако придется немного подождать, прежде чем смысл этого сравнения станет вам ясен.

НАИБОЛЕЕ часто используемое свойство атома — это то, что он имеет (или ему можно придать) значение. У одних атомов значение остается постоянным, другие являются переменными. Как и следовало ожидать, значение атома "1729" — целое число 1729 и оно остается постоянным. (Я провожу здесь различие между атомом, текстовое имя которого print name, или pname (цепочкой из четырех цифр "1729"), и его платоновой сущностью, которая заключается в том, что это число представляет собой сумму кубов двух других чисел, взятую двумя различными способами.) Значение атома nil также постоянно и равно nil! Еще один атом, значение которого постоянно и равно самому атому, это специальный атом "t".

За исключением атомов t, nil и атомов, имена которых являются числами, все атомы, вообще говоря, представляют собой переменные, т.е. им можно присвоить какие-то значения, а потом изменять эти значения произвольным образом. Как это делается на практике? Ну, например, можно присвоить значение 4 атому "пирог", для этого нужно напечатать джину Лиспа: "(setq пирог 4)". Однако это можно сделать и так: "(setq пирог (plus 2 2))" или даже: "(setq пирог (plus 1 1 1 1))". В любом из этих случаев сразу после того, как вы нажмете клавишу "возврат каретки", "пирог" получит значение 4, таким оно и останется, по крайней мере до тех пор, пока не будет выполнена

другая операция setq над атомом "пирог".

Лисп потерял бы свою привлекательность, если бы значениями атомов могли быть только числа. К счастью, значением атома может быть любой объект Лиспа — будь то какой-то другой атом или произвольный список. Например, значением атома "пи" можно сделать список, такой, как "(а б в)" или же "(plus 2 2)" вместо числа 4. Соответствие между атомом и его значением называют *связыванием* атома, другими словами, атом *связан* с этим значением. Чтобы установить новые связи для атомов "пирог" и "пи", мы снова воспользуемся операцией setq. В качестве иллюстрации ниже приведен короткий диалог с джином:

```
→ (setq пирог (plus 2 2))
4
→ (setq пи '(plus 2 2))
(plus 2 2)
→
```

Обратите внимание на то, сколь велико различие между значениями, присвоенными атомам "пи" и "пирог" в соответствии с нашими пожеланиями, формулировка которых отличается лишь присутствием или отсутствием маленького, но очень важного значка — одиночной кавычки перед внутренним списком "(plus 2 2)". В первом случае, когда кавычка отсутствует, внутреннее выражение "(plus 2 2))" должно *быть вычислено*. Возвращаемый результат — число 4, таково новое значение, присвоенное переменной "пирог". Во втором случае благодаря присутствию кавычки выражение "(plus 2 2)" не выполняется как команда, а рассматривается как пассивный элемент данных, просто как, например, кусок мяса, лежащего на прилавке в мясной лавке, — оно очень близко к "живому", но все же мертво. Таким образом, значение "пи" во втором случае — это список "(plus 2 2)", т.е. фрагмент программы на Лиспе. Следующий диалог с джином подтверждает правильность этой интерпретации значений атомов:

```
→ пирог
4
→ пи
(plus 2 2)
→ (eval пи)
4
→
```

Что означает оператор (eval пи)? Я хотел показать, как можно попросить джина *вычислить* значение выражения, вместо того чтобы просто *распечатать* это значение (eval — от evaluate, т.е. вычислить). Обычно система Лиспа автоматически производит вычисление одного уровня, однако при помощи операции eval можно вызвать вычисление второго уровня. (И конечно, повторяя эту операцию снова и снова, можно продолжать процесс последовательного вычисления сколь угодно долго.) Эта возможность в языке во многих ситуациях оказывается неоценимой, однако подробное обсуждение характера ее использования пока было бы преждевременным.

КАЖДЫЙ СПИСОК, за исключением nil, содержит по крайней мере один элемент. Первый элемент списка называется car. Таким образом, car от "(eval пи)" — это атом eval. Первые элементы списков "(plus 2 2)", "(setq x 17)", "(eval пи)" и "(car пи)" все являются названиями операций, или, как их чаще называют в Лиспе, *функций**. Однако car от списка вовсе не обязательно должен быть именем функции или даже атомом. Например, "((1) (2 2) (3 3 3))" — тоже список. Его car является списком "(1)", чей car в свою очередь — не имя функции, а просто число.

Если удалить первый элемент списка (его car), то что останется? Более короткий список. Он называется cdr от исходного списка, это слово произносится как "кадр" или "цедр". (Слова "car" и "cdr" остались в Лиспе на память от первой реализации Лиспа на машине IBM 704. Это сокращения от английских слов, означающих соответственно "содержимое адресной части регистра" и "содержимое уменьшаемой части регистра"; термины относились к специфическим особенностям аппаратуры, давно вышедшей из употребления.) В результате применения cdr к "(а б в)" мы получим список "(б в)", cdr от которого — это "(в)", а cdr от последнего — это nil. А nil уже не имеет cdr, точно так же как он не имеет car. Попытка взять car или cdr от nil приводит к появлению сообщения об ошибке (или по крайней мере должно привести к этому), так же как и попытка деления на нуль.

Приведем небольшую табличку, в которой показаны car и cdr от нескольких списков, просто для того, чтобы у вас не осталось никаких неясностей по этому поводу:

список	*car*	*cdr*
((а) б (в))	(а)	(б (в))
(plus 2 2)	plus	(2 2)
((car x) (car y))	(car x)	((car y))
(nil nil nil nil)	nil	(nil nil nil)
(nil)	nil	nil
nil	ERROR	ERROR

(Слово ERROR означает "ошибка".)

*По традиции функции Лиспа читаются как слитные слова. — *Прим. перев.*

В соответствии с тем, что car и cdr называют функциями, объекты, которыми они оперируют, называют аргументами. Таким образом, в команде "(plus пирог 2)" plus — это имя функции, а ее аргументы — это "пирог" и "2". При ее вычислении (так же, как и большинства других команд) джин сначала вычисляет значения аргументов и затем применяет функцию к этим значениям. Поскольку значение атома "пирог" равно 4, а значение атома "2" равно 2, то возвращается атом "6".

ПРЕДПОЛОЖИМ, что у вас есть список и вы хотели бы получить такой же список, но одним элементом больше. К примеру, пусть значением атома "x" является "(кекс печенье)" и вам нужно создать новый список "y", отличающийся от "x" тем, что он содержит еще один атом — скажем, "пирог" — впереди. В таком случае можно воспользоваться функцией cons (сокращение от construct — строить), действие которой заключается в создании нового списка из исходного и его нового элемента car. Вот как выглядит эта процедура:

```
→ (setq x '(кекс печенье))
(кекс печенье)
→ (setq y (cons 'пирог x))
(пирог кекс печенье)
→ x
(кекс печенье)
```

Тут следует отметить два момента. Я специально попросил распечатать значение x после выполнения операции cons, чтобы вы могли убедиться, что сам x не был изменен этой операцией. Она создала новый список и сделала его значением y, не трогая x. Другой примечательный момент состоит в том, что я опять воспользовался кавычкой перед атомом "пирог". А что было бы, если бы я ее не поставил? Результат был бы следующим:

```
→ (setq z (cons пирог x))
(4 кекс печенье)
```

Вспомним, что атом "пирог" все еще имеет значение 4, и, когда бы джин ни увидел атом без кавычки, содержащейся в команде, он воспользуется значением этого атома, а не его именем. (Всегда? Почти всегда. Очень скоро я объясню, в чем дело. А пока попробуйте сами найти исключение из этого правила — оно уже встречалось выше.)

Теперь несколько упражнений. Иногда в них будут некоторые нюансы. Следите за кавычкой! В этих упражнениях я пользуюсь функцией reverse, которая создает список, во всем подобный исходному, т.е. своему аргументу, но элементы в нем расположены в обратном порядке. Например, если приказать джину (reverse '((а б) (в г д))), то он

напечатает ((в г д) (а б)). Ответы джина на команды, перечисленные ниже, будут собраны вместе и приведены несколько позже.

→ (setq w (cons пирог '(cdr z)))
→ (setq v (cons 'пирог (cdr z)))
→ (setq u (reverse v))
→ (cdr (cdr u))
→ (car (cdr u))
→ (cons (car (cdr u)) u)
→ u
→ (reverse
'(cons (car u) (reverse (cdr u))))
→ (reverse
(cons (car u) (reverse (cdr u))))
→ u
→ (cons 'печенье
(cons 'кекс
(cons 'пирог nil)))

Приведем теперь ответы (напечатанные джином).

(4 cdr z)
(пирог кекс печенье)
(печенье кекс пирог)
(пирог)
кекс
(кекс печенье кекс пирог)
(печенье кекс пирог)
((reverse (cdr u)) (car u) cons)
(кекс пирог печенье)
(печенье кекс пирог)
(печенье кекс пирог)

В ПОСЛЕДНЕМ примере, где повторно применяется функция cons, продемонстрирован прием, который на жаргоне Лиспа называется "сконсить список". При этом вы начинаете с nil и затем раз за разом обращаетесь к функции cons. Эта процедура напоминает построение положительного целого числа (т.е. счет), когда, начиная с нуля, повторно применяется одна и та же операция (сложения). Однако, если в случае построения числа каждая последующая операция определена однозначно, в случае работы со списком у нас имеется бесконечное многообразие объектов, которые могут быть добавлены к списку, таким образом строится обширная разветвляющаяся структура, или дерево списков. Эта структура принципиально отличается от неразветвляющейся строки чисел. Именно поэтому я уподобил nil земле, где коренятся все структуры — они похожи на "деревья", растущие из атома nil.

Как уже говорилось выше, джин *не всегда* заменяет атомы (перед которыми нет кавычки) их значениями. Существуют случаи, когда функция рассматривает свои аргументы как имеющие кавычку, хотя на самом деле они ее не имеют. Наверное, вы уже вернулись к предыдущим примерам и нашли такой случай? Это было нетрудно сделать. Речь идет о функции setq. При выполнении этой команды первый атом берется как он есть, его значение не вычисляется. На самом деле буква "q" в слове setq — это сокращение от слова quote (т.е. цитировать, брать в кавычки); благодаря присутствию этой буквы первый аргумент рассматривается как имеющий кавычку. Ситуация несколько усложнится, когда вы познакомитесь с функцией set, во всем подобной setq, за исключением того, что она вычисляет свой первый аргумент. Поэтому если значением атома "x" является атом "k", то выполнение команды "(set x 7)" ничего не сделает с "x" (его значение останется равным атому "k"), но значение атома "k" станет равным 7. Посмотрите внимательно на следующие примеры:

→ (setq a 'b)
→ (setq b 'c)
→ (setq c 'a)
→ (set a c)
→ (set c b)

А теперь скажите: какие значения имеют атомы "a", "b" и "c"? Не подглядывайте, сейчас я дам ответ. Эти значения равны соответственно "a", "a" и "a". Тут нетрудно запутаться. Однако в утешение я могу вам сказать, что функцией set пользуются в Лиспе не так уж часто и подобные затруднения возникают редко.

С ПСИХОЛОГИЧЕСКОЙ точки зрения одна из наиболее интересных возможностей программирования заключается в способности определять новые составные операции на базе старых. Повторяя эту процедуру снова и снова, можно создать огромный репертуар все более сложных операций. Это несколько напоминает ранние стадии биологической эволюции, когда все более сложные молекулы возникали из менее сложных по восходящей спирали. Можно также провести сравнение с промышленной революцией, когда созданные на более ранних стадиях простые машины помогали людям строить более сложные машины, которые в свою очередь использовались для создания еще более сложных машин, и так далее по восходящей спирали сложности и технических возможностей. На каждой новой стадии, будь то эволюция видов или промышленная революция, конструкции становились более гибкими и замысловатыми, более "разумными", но зато и более уязвимыми в отношении скрытых, "тонких" нарушений или поломок.

Аналогичная ситуация наблюдается при программировании на Лиспе, только здесь мы имеем дело не с молекулами или машинами, а с функциями Лиспа, определенными на базе уже известных функций языка. Предположим, например, что вам понадобилась функция, всегда возвращающая последний элемент списка, точно так же как car возвращает первый элемент. Среди стандартных функций Лиспа такой функции нет, но вы можете ее легко создать сами. Знаете как? Чтобы получить последний элемент списка с именем "спис", вы можете просто обратить порядок элементов "спис" и взять первый элемент нового списка: "(car (reverse спис))". Дадим этой операции имя rac (car в обратном порядке), воспользовавшись функцией

→ (def rac (lambda (спис)
(car (reverse спис))))

Функция def применяется для того, чтобы *определить новую функцию*. Слово lambda, за которым следует "(спис)", означает, что определяемая функция имеет один параметр, называемый формальным параметром, поскольку он является абстрактной, фиктивной переменной. Этот параметр был назван нами именем "спис". (Его можно было назвать как угодно.) Вообще список параметров должен непосредственно следовать за словом lambda. После того как желание def было выполнено, функция rac становится так же хорошо знакомой джину, как и функция car. Таким образом, (rac '(за руку цап))" возвращает атом "цап". Функцией rac можно теперь воспользоваться при определении других новых функций. Эти определения растут как снежный ком, и через некоторое время вы будете ошеломлены богатством возможностей, которыми располагаете.

Приведем простой пример. Предположим, вам предстоит иметь дело с большим количеством длиннющих списков и вы хотели бы завести на каждый такой длинный список один коротенький, состоящий лишь из первого и последнего элементов длинного списка. Для этого достаточно определить функцию с одним параметром:

(def краткое-содержание
(lambda (длинный-список)
(cons (car длинный-список)
(cons (rac длинный-список) nil))))

Следовательно, если применить функцию "краткое содержание" к полному тексту художественного произведения, например большого романа (рассматривая его как длинный список слов), то получим более короткий список "(Ну зависимость)". К сожалению, повторное применение этой функции к новому списку уже ничего не упростит.

Было бы полезно иметь в своем распоряжении функцию, обратную функции "краткое содержание", которая по двум заданным словам могла бы воссоздать весь текст романа, начав и кон-

чив его этими словами. Назовем эту функцию "писатель". Так, выполнение оператора "(писатель Все нее)" приведет к тому, что джин Лиспа воссоздаст весь текст романа. Определение этой функции мы оставляем читателю в качестве упражнения.

МОЖЕТ показаться, что главная цель, стоящая перед Лиспом и языками, подобными ему, цель, к которой нужно стремиться и которая достижима, должна заключаться в том, чтобы (1) каждый оператор возвращал *значение* и (2) чтобы через это возвращенное значение, и *только* через него, достигалось бы действие оператора. Первая мысль означает, что значения передаются "наверх" от самых внутренних обращений к функциям к внешним обращениям, пока результирующее значение всего оператора не будет возвращено в качестве ответа. Вторая мысль означает, что в продолжение всех этих обращений ни один атом не меняет своего значения (кроме тех атомов, которые играют роль формального параметра). Требование (1) выполняется во всех известных мне диалектах Лиспа, но требование (2) в некоторых ситуациях не выполняется.

Например, если x связан с "(а б в г д)" и вы задаете оператор "(car (cdr (reverse x)))", то сначала будет вычислено значение "(reverse x)", затем это значение будет передано "наверх" функции cdr, которая вычислит cdr от полученного списка, и, наконец, уже более короткий список передается функции car. Она выделяет из него один элемент, атом "г", и возвращает его в качестве результата. Тем временем атом "x" не претерпел никаких изменений — он попрежнему связан со списком "(а б в г д)".

Казалось бы, выражение "(reverse x)" должно было изменить значение x, обратив порядок его элементов, например, так же, как устная команда "выверни свой свитер наизнанку" подействовала бы на свитер. На самом деле, выполнение желания "(reverse x)" не в большей степени повлияет на значение x, чем выполнение желания "(plus 2 2)" может повлиять на значение 2. Вместо этого выполнение "(reverse x)" приводит к созданию *нового* безымянного списка, такого же, как x, но с обратным порядком элементов. Этот список является вычисленным значением оператора, или результатом, который возвращает оператор. Само же по себе значение x не изменяется. Аналогичным образом вычисление "(cons 5 пи)" не изменит значение списка с именем "пи", оно просто *возвращает* новый список, car от которого — 5, а значение "пи", каким бы оно ни было, — его cdr.

Это резко отличается от поведения тех функций, которые оставляют после своего выполнения "побочные эффекты". Чаще всего такие эффекты выражаются в виде измененных связей переменных, хотя есть и другие возможности, например в виде ввода или вывода данных. Наиболее типичной "вредной" командой является setq, и сторонники "аппликативного" стиля программирования — школа, запрещающая какие бы то ни было побочные эффекты — очень расстраиваются даже при простом упоминании о функции setq. С их точки зрения, результаты должны получаться исключительно путем вычисления функциями значений и последующей передачи их другим функциям.

Лишь один вид связывания допускается сторонниками такого стиля — временные "лямбда-связывания", устанавливаемые, когда функция применяется к своим аргументам. Всякий раз, когда происходит обращение к функции, ее формальные параметры получают временные "лямбда-связи". Эти связи устанавливаются аналогично тем, которые порождает setq, но они временные. Когда данная функция завершает свои вычисления, они пропадают. Например, при вычислении "(rac '(а б в))" формальный параметр этой функции "спис" связывается со списком "(а б в)", т.е. происходит лямбда-связывание, однако, как только ответ "в" передан программисту или функции, запросившей выполнение rac, значение атома "спис", с которым работала функция rac, оказывается забытым. И если вы запросите значение атома "спис", интерпретатор Лиспа ответит вам, что "спис" является "несвязанным" атомом. Программисты аппликативной школы гораздо охотнее пользуются лямбда-связыванием, чем операцией setq.

Я не являюсь фанатичным противником setq и других функций, вызывающих побочные эффекты; хотя аппликативный стиль элегантный, он представляется мне непрактичным, коль скоро речь идет о построении больших программ искусственного интеллекта. Поэтому я не рекомендую вам всегда придерживаться аппликативного стиля, хотя по возможности я сам стараюсь его придерживаться. Строго говоря, при таком программировании вы даже не можете определить новые функции, поскольку оператор def приносит неисчезающее изменение в память джина — определение новой функции навсегда остается в памяти. Таким образом, согласно чисто аппликативному подходу, определения функций, подобно связываниям переменных, должны быть лишь временными, и сразу после использования их надо отбрасывать. Это, конечно, крайний "аппликативизм".

Чтобы поупражняться, рассмотрим еще несколько простых определений функций:

```
→ (def rdc
      (lambda (спис)
      (reverse (cdr (reverse спис)))))
→ (def snoc
      (lambda (x спис)
      (reverse
      (cons x (reverse спис)))))
→ (def дважды (lambda (n) (plus n n)))
```

Функции rdc и snoc аналогичны cdr и cons, но они действуют в обратном порядке. Следовательно, rdc от "(а б в г д)" дает "(а б в г)", а в ответ на "(snoc 5 '(1 2 3 4))" вы получите "(1 2 3 4 5)".

ВСЕ ЭТО, конечно, в какой-то степени интересно, но если вы хотите увидеть, как джин делает *действительно* удивительные вещи, то ему нужно позволить принимать решения, основанные на событиях, которые уже произошли до этого. Такие желания называют *условными.* Вот типичный пример:

```
→ (cond ((eq x 1) 'земля)
           ((eq x 2) 'море))
```

Значение, возвращаемое этим оператором, будет либо атомом "земля", если x имеет значение 1, либо атомом "море", если значение x равно 2. Во всех других случаях возвращаемым значением будет nil (например, если x равен 5). Атом eq — имя часто используемой функции Лиспа, которая возвращает атом "t" (от true — истинный), когда ее аргументы имеют одно и то же значение, и атом nil, когда эти значения не одинаковы.

Оператор cond — это список, car от которого является именем этой функции "cond", за ним следует произвольное количество предложений, или *cond-предложений*, каждое из которых представляет собой двухэлементный список. Первый элемент — это "условие", а второй элемент — "результат". Условия одно за другим проверяются джином Лиспа в том порядке, в котором они стоят. Как только джин находит предложение, условие которого оказывается выполненным, или истинным (уточним: условие, которое возвращает что угодно, но не nil), он приступает к вычислению результата, содержащегося в этом предложении, и вычисленное значение возвращается в качестве значения всего оператора cond. Ни на одно из последующих предложений джин даже не посмотрит. Возможно, это звучит сложнее, чем оно есть на самом деле. Можно просто сказать, что джин отыскивает первое попавшееся условие, которое удовлетворяется, и возвращает результат, соответствующий этому условию.

Часто возникает необходимость во "всеулавливающем" условии в конце условного оператора, которое наверняка выполняется, так что, даже если все остальные условия не сработают,

по крайней мере это последнее окажется выполненным и вместо nil будет возвращен соответствующий этому условию результат. Придумать условие, значение которого не равно nil, очень просто — примем его равным "t", как в следующем примере:

```
→ (cond ((eq x 1) 'земля)
        (eq x 2) 'море)
        (t 'воздух))
```

В зависимости от того, чему равно значение x, мы получим в качестве значения оператора "землю", "море" или "воздух", но мы никогда не получим nil. Вот еще несколько операторов cond, попробуйте сыграть роль джина:

```
→ (cond ((eq (eval пи) пирог)
      (eval (snoc пирог пи)))
      (t (eval (snoc (rac пи) пи))))
→ (cond ((eq 2 2) 2) ((eq 3 3 (3))
→ (cond (nil 'нет-нет-нет)
        ((eq '(car nil) '(cdr nil))
        ('гммм)
        (t 'да-да-да))
```

Ответы: 8, 2 и "да-да-да". Вы заметили, что перед "(car nil)" и "(cdr nil)" стоит кавычка?

В ЗАКЛЮЧЕНИЕ этой статьи я покажу вам некое семейство определений функций, обладающих настолько очевидной закономерностью, что, казалось бы, джин "сообразит, в чем дело», увидев первые примеры. К сожалению, джин Лиспа на удивление туп (или по крайней мере прикидывается тупым), и он не сделает никаких выводов, пока они не будут ему указаны в явном виде. Взгляните на это семейство:

```
→ (def квадрат
      (lambda (k) (times k k)))
→ (def куб
      (lambda (k) (times k (квадрат k))))
→ (def 4-я-степень
      (lambda (k) (times k (куб k))))
→ (def 5-я-степень
      (lambda (k)
      (times k (4-я-степень k))))
→ (def 6-я-степень
      (lambda (k)
      (times (5-я степень k))))
```

Подумайте над таким вопросом: можете ли вы изобрести определение функции с *двумя* параметрами, которая "одним ударом" заменила бы всю цепочку определений, приведенных выше? (Ответ мне известен, и я сообщу его вам в следующий раз.) Более конкретно вопрос заключается в следующем. Как определить функцию с двумя параметрами по имени "степень", такую, чтобы, например, результат вычисления "(степень 9 3)" был равен 729, а "(степень 7 4)" — соответственно 2401? Вы уже располагаете всеми необходимыми средствами, нужно лишь проявить немного изобретательности.

НАКОНЕЦ, мне хотелось бы поделиться с вами новостью о недавно открытом животном глазункианском дикогрызе, названном так потому, что он водится только на острове Глазункия (на который претендует Верхний Битбо, хотя остров находится вплотную к берегам Бурронемеды). На что похож дикогрыз? Это — своеобразная разновидность дикобраза, иглы которого — у него их почему-то всегда девять (во Внешней Глазункии) или семь (во Внутренней Глазункии) — сами представляют собой более мелких дикогрызов. Ого! Определенно может показаться, что здесь имеет место бесконечная регрессия! Но это не так. Я просто забыл сказать, что существуют дикогрызы наименьшего размера — нульдюймовые, которые, как это ни странно, совершенно лысые, у них нет ни одной иглы. Поэтому, к счастью (или, может быть, к сожалению — это зависит от вашей точки зрения), угроза бесконечной регрессии отпадает.

Читателям, изучающим зоологию, возможно, интересно будет узнать, что иглы пятидюймовых дикогрызов являются всегда четырехдюймовыми дикогрызами и т.д. А читатели, увлекающиеся антропологией, по-видимому, будут не менее заинтригованы, узнав, что жители Глазункии (как Внешней, так и Внутренней) пользуются носами (да, именно носами) нульдюймовых дикогрызов как меновой единицей — странно, не правда ли? Однако не нам с вами судить о древних обычаях Внешней и Внутренней Глазункии! Стало быть, поскольку крупный дикогрыз — скажем, трех или четырехдюймовик — имеет много-много таких крошечных носиков, он представляет собой большую ценность. Ценность дикогрыза иногда называется "покупательной степенью" или, короче, просто "степенью". Например, двухдюймовик, найденный во Внутренней Глазункии, почти вдвое ценнее двухдюймовика, найденного во Внешней Глазункии. Или я перепутал, и все наоборот?

Во всяком случае, зачем я все это рассказываю? Просто мне показалось, что это вам будет интересно. Кто знает? Возможно, в один прекрасный день вам придется посетить Глазункию (Внутреннюю или Внешнюю), и тогда все это может пригодиться.

Книги

Альберт Эйнштейн — ученый и человек
Часы внутри нас
Мир аммонитов

ФИЛИП МОРРИСОН

Абрахам Пайс. БОГ ИЗОЩРЕН...: АЛЬБЕРТ ЭЙНШТЕЙН — УЧЕНЫЙ И ЧЕЛОВЕК.
SUBTLE IS THE LORD...: THE SCIENCE AND THE LIFE OF ALBERT EINSTEIN, by Abraham Pais. Oxford University Press ($ 25)

В 1921 г. Альберт Эйнштейн впервые посетил Соединенные Штаты. В Принстоне он узнал, что на обсерватории Маунт-Вилсон проведены новые эксперименты, доказывающие увлечение эфира. Реакцией Эйнштейна на это сообщение стала известная теперь фраза, высеченная позднее на камне в Принстоне: "Raffiniert ist der Herr Gott, aber boshaft ist Er nicht" — «Бог изощрен, но не злонамерен»*. Эйнштейн перефразировал свое высказывание, когда в 1930 г. согласился на его увековечение: «Природа хранит свои тайны с присущим ей благородством, не прибегая к мошенническим уловкам». Само понятие эфира вскоре отошло в прошлое. Современные представления о вакууме как о сложном пространстве с виртуальным заполнением, влияющем и на энергию частиц, и на расширение Вселенной, «не имеют ничего общего с эфиром... с которым покончено раз и навсегда».

Эйнштейн был великим мастером изложения своих идей и отличался особым даром ярко и необычайно точно выражать свою мысль. Вопреки сложившемуся мнению, утверждает автор книги, Эйнштейн был прекрасным психологом: оценки, которые он давал своим современникам, лучше всего опровергают суждения о нем как о простодушном человеке. В книге с глубокой симпатией, но вполне объективно рассказывается о его жизни и работе в области физики, которой была отдана вся его жизнь. Материалом для исследования, проведенного Пайсом, послужили работы самого Эйнштейна, дополненные сведениями из многочисленных источников.

Среди них неопубликованные «Дополнения к биографии брата», написанные младшей сестрой Эйнштейна Майей в 1924 г., новые книги, вышедшие в ФРГ, Швейцарии и даже Японии, неопубликованная рукопись из библиотеки Моргана, письма из хранилищ Цюриха, Мюнхена и Западного Берлина и, разумеется, личный архив Альберта Эйнштейна. Великолепным гидом по материалам этого колоссального архива была Элен Дюкас, секретарь и незаменимый помощник Эйнштейна (скончавшаяся в 1982 г.). При знакомстве с книгой прежде всего бросается в глаза, насколько органично в повествовании о физике, умершем почти три десятилетия назад, вплетаются ссылки на новейшие научные статьи из журнала "Physical Review Letters". Нетрудно догадаться, что автор книги — крупный ученый-физик.

Профессор Пайс, известный в своей области теоретик старшего поколения, в последние годы проявляет интерес к истории физики. Результатом этого увлечения и явилась настоящая книга — дань любви и уважения к Альберту Эйнштейну. Автору посчастливилось в послевоенные годы неоднократно беседовать с Эйнштейном как во время их совместных прогулок по Принстону, так и в Институте высших исследований. Необыкновенно интересны краткие зарисовки встреч старика Эйнштейна с молодым голландским физиком, специалистом по квантовой теории поля, только что приехавшим из Института Нильса Бора в Копенгагене. Так, например, автор рассказывает: «Эйнштейн внезапно остановился, обернулся ко мне и спросил, верю ли я, что Луна существует лишь в тот момент, когда я на нее смотрю».

Последняя встреча состоялась за несколько месяцев до кончины Эйнштейна. Семидесятипятилетний ученый был болен и некоторое время не появлялся в институте. Пайс, собиравшийся в длительный отпуск, зашел к нему домой попрощаться. «Он работал. Как только я вошел, он отложил свой блокнот... Мы провели за приятной беседой около получаса... Я сказал, что мне пора идти... У самой двери я оглянулся. Он сидел в кресле с блокнотом на коленях и карандашом в руке... Он снова работал».

Наиболее удачны разделы книги, где подробно говорится о столь горячо любимой Эйнштейном работе и где [в терминах ixerei (иксов)] на лаконичном языке физиков-теоретиков дается ее тщательный анализ, неизменно сопровождаемый, однако, текстом и пояснениями. Читатель, хоть немного подготовленный к восприятию такой формы изложения, получит, разумеется, от книги максимальную пользу. Однако и тот, кто не разбирается в формулах, найдет материал для расширения своего кругозора и почерпнет новые весьма интересные сведения о жизни того, кто создал эти уравнения. Из всех книг об Эйнштейне именно в книге Пайса наибольшее внимание уделено его научным исследованиям, их истокам, неожиданным поворотам мысли, а также развитию идей Эйнштейна другими исследователями.

Несмотря на то что многие события общественной и личной жизни Эйнштейна давно уже стали достоянием широкой публики, читатель найдет в книге ряд ранее не известных фактов, а также оригинальную и убедительную интерпретацию старых. Мальчик родился с головой необычной формы ("Viel zu dick" — «слишком большая», как сказала бабушка) и очень поздно начал говорить. Однако, как оказалось, учился он совсем не плохо. Его мать писала о своем семилетнем сыне: «Альберт опять первый ученик в классе. Он принес отличный аттестат». За годы учебы в гимназии Луитпольда мальчик неизменно был одним из первых учеников не только по математике, но и по латыни. Миф о плохом ученике оказывается необоснованным. Правда, Эйнштейн, отличавшийся независимой манерой поведения, сам рассказывал, что в старших классах один из преподавателей, бывало, возражал против его присутствия на уроках: «Вы сидите в последнем ряду и ухмыляетесь и, следовательно, не проявляете должного уважения к учителю».

Эксперименты Эйнштейна и патенты, полученные им вместе с талантливыми соавторами, зачастую по инициативе последних, составляют довольно большой список — от первой успешной демонстрации вращения проводящей жидкости, возникающего в магнитном поле под действием электрического тока, до электромагнитного насоса в бытовых холодильниках. В наши дни подобные насосы используются в ядерных реакторах.

Эйнштейн был женат дважды, первый раз на своей однокурснице, второй — на преданной и поматерински заботливой кузине, однако образцовым мужем и отцом он никогда не был даже в собственных

* Это изречение Эйнштейна высечено над камином общей гостиной в Файн-Холле, здании Института математики и теоретической физики Принстонского университета. — *Прим. перев.*

глазах. Когда умер его друг Бессо, Эйнштейн, переживший его лишь на месяц, писал: «Больше всего меня восхищает в этом человеке то, что ему удалось прожить с женщиной много лет не просто в мире, но и в полном согласии — в чем я, надо признаться, дважды потерпел почти полное фиаско».

Необыкновенно увлекательно следить за развитием его научной мысли. Пайс отмечает сходство между ранними научными успехами Эйнштейна и блестящей карьерой Моцарта: та же легкость, свежесть, лавина успеха. Дальнейшая его деятельность развивается скорее в бетховенском ключе — упорство и труд, выковывающие наконец мастера, которому нет равных. Общая теория относительности и основанная на ней современная космология были четко сформулированы к 1918 г. Из скрупулезного анализа, проделанного Пайсом, можно сделать следующий вывод: долгое странствие, в которое физик-одиночка Эйнштейн пустился после 1925 г. и которое продолжалось до конца его жизни, — попытка создания единой теории поля — не было вызвано только стремлением объединить разные виды сил в рамках одной теории. Актуальной эта проблема стала в физике наших дней. Его усилия были направлены на то, чтобы выделить квантовую зернистость из мира непрерывного поля, сделать частицы и вероятность следствиями инвариантности нелинейных, но строго причинных полей, используемых в теории гравитации. Не все теоретики разделяют энтузиазм Эйнштейна, стремившегося вернуть ньютоновскую причинную связь миру, надежно подчиняющемуся, с точностью до многих десятичных знаков, хитроумным законам квантованных полей.

Знакомясь с открытиями, сделанными Эйнштейном, невольно задумываешься, какой была бы современная физика без этого плодовитейшего теоретика нашего века. Его первая выдающаяся работа о реальности молекул была отчасти предвосхищена Уиллардом Гиббсом и Марианом Смолуховским, независимо получившими сходные результаты. К открытию специальной теории относительности был близок Анри Пуанкаре, который понял необходимость формальных преобразований и переосмысления понятия одновременности, но не сумел объединить одно с другим. Мог ли появиться другой молодой ученый, который сумел бы это сделать? Нам остается только гадать.

Создание теории гравитации требовало еще более мощной интуиции. Известно, однако, что к формулированию этой теории была близка и группа геттингенских математиков во главе с Давидом Гильбертом. Гильберт самостоятельно вывел уравнения Эйнштейна, правда уже ознакомившись с некоторыми идеями более ранних работ последнего. Быть может, лишь фотон — эта волна-частица максвелловского электромагнитного поля — для своего открытия более всего нуждался именно в Эйнштейне. Начиная с работы по фотоэффекту в 1905 г. и далее в целой серии блестящих идей непреходящей ценности об излучении и веществе, Эйнштейн неизменно защищал реальность фотона. Даже Бор однажды чуть не отрекся от фотона: в 1924 г. он усомнился в сохранении энергии и импульса в элементарных атомных процессах.

Эйнштейн был непоколебим: еще раньше, в 1909 г., он писал: «Я считаю, что на следующем этапе будет создана теория света, объясняющая единство волновых и корпускулярных свойств... (которые) не следует рассматривать как исключающие друг друга». Однако, когда такая теория была сформулирована (около 1930 г.), Эйнштейн, несмотря на целый ряд экспериментальных подтверждений, не принял ее, считая ошибочной и не ведущей в конечном счете к истине. Пайс пытается найти оправдание такой философской позиции Эйнштейна. Действительно, как могло случиться, что ее занял ученый, от которого человечество восприняло не только представление о новой структуре пространства-времени, но и реальность сосуществования волновых и корпускулярных свойств? Сам Эйнштейн однажды сказал: «Первооткрывателю все созданное его воображением кажется столь необходимым и естественным, что он склонен относиться к своим творениям как к чему-то реально существующему». Дважды ему удавалось проникнуть в глубины мироздания, строго придерживаясь исходных посылок, выбранных из соображений простоты. В третий раз этому не суждено было случиться.

«Эйнштейн был самым свободным человеком из всех, кого я знал... властелином своей судьбы», смело бросающим вызов природе и бесстрашным перед лицом времени и смерти. Не было ничего трагического ни в его отношении к квантовой теории, ни в его безуспешных поисках единой теории поля. Так заканчивает свою книгу Пайс, ученый и друг Эйнштейна.

Мартин К. Мур-Ид, Фрэнк М. Зульцман и Чарлз А. Фуллер. ЧАСЫ ВНУТРИ НАС: ФИЗИОЛОГИЯ ЦИРКАДНОЙ СИСТЕМЫ СИНХРОНИЗАЦИИ

THE CLOCKS THAT TIME US: PHYSIOLOGY OF THE CIRCADIAN TIMING SYSTEM, by Martin C. Moore-Ede, Frank M. Sulzman and Charles A. Fuller. Harvard University Press ($ 25)

КОШКА, которая гуляла сама по себе, бродила, где ей вздумается. Удивительнее, что, как утверждают факты, кошки и ко *времени* относятся, как им вздумается. За кошкой наблюдали в течение месяца, ее активность расписали по часам и не обнаружили почти никакой зависимости от времени. Периоды активности и отдыха сменяют друг друга нерегулярно, с промежутками в один-два часа, более длительные внутренние ритмы не обнаруживаются, нет даже зависимости от времени суток. Мыши и люди в этом отношении на кошек не похожи. Для них характерны четкие ритмы активности. Если человека или мышь поместить в тщательно изолированную лабораторию и лишить возможности как-либо определять время, наблюдаются спонтанные циркадные ритмы, близкие, но не совпадающие с 24-часовым периодом вращения Земли. За исключением кошек, собак и рысей, у большинства млекопитающих — да и у других живых организмов — есть явные внутренние циркадные ритмы. Период цикла характерен для вида, хотя у разных особей одного вида он несколько различается. Мыши, как правило, «спешат» — их стабильный внутренний ритм составляет около 23,5 часов. Люди немного «отстают»: полный оборот стрелок наших внутренних часов совершается примерно за 25 часов.

В этой богатой фактическим материалом книге доходчиво и не слишком специально описывается интереснейший аспект физиологии человека: система внутренних ритмов. Она тесно связана со сном и бодрствованием; безусловно, систему внутренних ритмов следует считать проявлением все еще не исследованных до конца многообразных неосознаваемых связей между телом и мозгом. Материал изложен ясно и последовательно, математических формул лишь необходимый минимум, зато много графически представленных экспериментальных данных. Авторы стремились сделать какие-то практические выводы, которые пригодились бы в быту и в медицине.

Мы живем по часам, пользуясь циферблатом или ориентируясь по вечным естественным часам — Солнцу. Многие особенности организма, не говоря уже о ритме сна и бодрствования, зависят от смены дня и ночи. Как во внутренних процессах, так и во взаимодействии организма с внешней средой есть отчетливая суточная ритмичность. Например, показано, что у лабораторных мышей 50 — 60 физиологических функций циклически изменяется в зависимости от времени суток.

На одном из графиков изображены суточные ритмы человека. Графики ритмических изменений различных параметров очень различны по форме, амплитуде и фазе. Ночью мы спим, в это время уровень гормона роста в крови максимален, а внутренняя температура тела и выведение калия с мочой, наоборот, минимальны. Уровень кортизола минимален во время глубокого сна, но незадолго до пробуждения резко возрастает и достигает максимальной величины. Врач, наблюдающий ночью за состоянием тяжелобольного, должен иметь в виду такие колебания физиологических и биохимических показателей. Ритмы изменяются с возрастом: примерно до месяца младенцы спят и просыпаются почти так же беспорядочно, как кошки (чередование сна и бодрствования у младенца зависит, по-видимому, и от сроков кормления). Затем возникает отчетливый ритм, который немного отстает от смены дня и ночи, а примерно с четвертого месяца жизни синхронно с чередованием времен суток, как бы подстраиваясь к этому внешнему ритму, начинают работать внутренние часы.

При внимательном изучении такого подстраивания выявляются новые особенности естественных внутренних ритмов. Внешний водитель ритма может в некоторых пределах изменять внутренний ритм — обычно на несколько часов в ту или другую сторону. Чтобы навязать новый ритм, достаточно совсем слабого раздражителя: например, с помощью короткой вспышки света один-единственный раз в сутки можно «синхронизировать» беличью обезьяну (эта маленькая обезьянка, как и человек, ведет дневной образ жизни). Замечательно, что, если устранить внешний водитель ритма, восстанавливается прежний внутренний ритм. Такие же данные есть и в отношении человека.

В одном из типичных опытов у испытуемого в полностью изолированной лаборатории в течение первых пяти суток сохранялся четкий 24-часовой ритм, синхронный с вращением Земли. Затем ритм сменился другим, 25-часовым. Этот ритм оказался устойчивым, и ему следовали как смена сна и бодрствования, так и кривая температуры тела. Примерно через пять недель в организме произошли какие-то перемены: цикл изменения температуры стал короче, хотя все еще немного превышал 24 часа, а цикл смены сна и бодрствования значительно удлинился. В нем правильно, но довольно прихотливо чередовались долгие и короткие периоды сна. Внутренние «сутки» удлинились до 30 с лишним часов.

Очевидно, в нас действует не одна, а две системы ритмов. Их взаимодействие хорошо описывается простой математической моделью. Одни часы регулируют температуру тела и связанные с ней функции, а другие — смену сна и бодрствования и функции, имеющие отношение к сну. Есть данные о локализации этих двух систем (для человека пока что таких данных нет). Первые часы, видимо, связаны с имеющимся у большинства млекопитающих органом — парной группой особых нейронов гипоталамуса. Многочисленные операции на обезьянах позволили определить точное местоположение этого участка. Врачам давно известно, что у человека при опухоли мозга вблизи аналогичной структуры может серьезно нарушаться ритм чередования сна и бодрствования. Еще более поразительно, что, если обезьяне впрыснуть в сетчатку глаза радиоактивный индикатор и сразу после этого провести аутопсию, метка обнаруживается в нервных волокнах, идущих от сетчатки прямо к «часам» в гипоталамусе. Пожалуй, теперь понятно, отчего мы просыпаемся на рассвете.

О вторых часах мы знаем меньше. Есть данные, что они также расположены в гипоталамусе, поблизости от первых, но отдельно от них. Связанные с сетчаткой глаза часы регулируют температуру тела, содержание гормона роста в крови и ту фазу сна, которая протекает без сновидений. Собственный ритм этих часов очень медленный, но на смену дня и ночи он обычно реагирует. Вторые часы, с 25-часовым ритмом, управляют фазой сна со сновидениями, выделением калия с мочой и теплообразованием в тканях тела. Их ритм передается на первый, медленный осциллятор, который посылает обратно сигналы, несущие информацию об освещенности и ее изменениях. Вместе эти две системы и создают наш привычный 24-часовой ритм. Разумеется, в книге все излагается гораздо подробнее. Авторы книги — опытные экспериментаторы: физиолог, биолог и невролог. В течение семи лет они вели плодотворную совместную работу, изучая циркадные ритмы.

С тех пор как человек овладел огнем, он становился все более независимым от ритма вращения Земли. Сдвиг времени в летние месяцы, посменная работа, перелеты на самолетах из одного часового пояса в другой — все это сильно влияет на древнюю систему циркадных ритмов. Стоит упомянуть и о жителях Арктики, на долгие месяцы лишенных солнца, которое всегда служило сигналом. А ведь родина *Homo sapiens* — тропические широты, и наши предки привыкли к восходам и заходам солнца. Случаются страшные истории: как-то раз пилот пассажирского самолета заснул в момент приземления, когда до посадочной полосы оставалось всего 60 м (к счастью, второй пилот не спал). В 1978 г. реактивный лайнер пролетел над аэродромом Лос-Анджелеса и приземлился на автопилоте более чем в 160 км от него — весь экипаж спал! Реактор на острове Тримайл обслуживал персонал, работавший посменно, причем время дежурства сдвигалось на несколько часов раз в неделю. Это было удобно, но люди не успевали приспособиться к новому режиму. Во флоте считают, что для команды ракетной подводной лодки 8-часовые смены — чересчур большая нагрузка. Однако, поскольку экипажа на борту хватает только на три вахты, «сутки» сокращены до 18 часов, что идет вразрез с требованиями физиологии, и люди испытывают сильный стресс.

Путешественникам лучше не менять режим, если они приехали в другой часовой пояс ненадолго. Если же перестроиться необходимо, предпочтительнее «отставать» по фазе. Гораздо легче приучить нашу медленную систему двойных часов к дням, которые на два часа длиннее, чем к дням, укороченным на те же два часа. Опережение времени при движении на запад переносится легче, чем отставание при перемещении на восток. Некоторые виды бессонницы, связанные со сменой часового пояса, легко поддавались лечению тщательным соблюдением режима с 27-часовыми «сутками». Изучение циркадных ритмов сейчас находится в захватывающе интересной стадии. Физиология человека переходит в совершенно новое измерение.

Ульрих Леман. АММОНИТЫ: ИХ ЖИЗНЬ И МИР.

THE AMMONITES: THEIR LIFE AND THEIR WORLD, by Ulrich Lehmann. Cambridge University Press ($ 19.95)

В ГЛУБИНАХ морей от островов Фиджи до Филиппин охотится по ночам жемчужный кораблик, наутилус. Медленно скользит он во тьме, и струи воды развевают венец тонких щупалец. Наутилус — живое ископаемое, как и целакант. Он почти такой же, как его предки, аммониты, жившие в ордовикский период. Задолго до появления первых рыб эти существа уже владели морями. В те времена многокамерные раковины не у всех аммонитов были спиральными.

В континентальных известняках на юго-западе Германии встречается множество ископаемых раковин, очень похожих на изящные плоские спиральные раковины наутилуса. Раковины одной из древних форм аммонитов так распространены, что их изображение даже вошло в немецкую геральдику. За жемчужный блеск и изысканную форму раковины аммониты заслуженно считают одним из самых красивых ископаемых на свете. Порой в раковинах сохраняется сама известковая основа.

Но окаменелые раковины — редкость, чаще попадаются слепки; они бывают такие четкие, что видны все детали внутреннего строения.

Ископаемые остатки аммонитов уже давно коллекционируют и изучают. В пещере близ Ульма есть изображение аммонита, сделанное палеолитическим человеком.

В настоящее время известно несколько тысяч форм размером от долей миллиметра до двух метров в поперечнике. Остатки их в изобилии встречаются по всему миру. Эпоха аммонитов охватывает более ста миллионов лет. Они исчезли одновременно с динозаврами в конце мелового — начале третичного периода. Уже лет сто по деталям строения раковин аммонитов определяют возраст геологических слоев, но биологией самих животных пренебрегали, рассматривая их всего лишь как руководящие ископаемые. Собственно говоря, мы очень мало знаем даже о современных наутилусах, их стали изучать в аквариумах только в последнее время.

Книга профессора Лемана — полное, прекрасно иллюстрированное введение в биологию головоногих, строивших спиральные раковины. Леман — палеонтолог, он внес существенный вклад в формирование новых концепций в этой области. Несмотря на то что предварительной специальной подготовки от читателя не требуется, сложность и разнообразие форм, да и необычность самой жизни морских моллюсков делают книгу несколько трудной для неспециалиста. Зато новые гипотезы, основанные на последних находках, захватывающе интересны. Эта небольшая по объему книга сторицей вознаградит того, кто, увлеченный загадками палеонтологии, не пожалеет труда, чтобы во всем разобраться.

Ключом к пониманию образа жизни аммонитов послужили находки их челюстей. Долгое время полагали, что у аммонитов вообще не было челюстей, то есть не было ничего подобного сильному, острому, похожему на клюв попугая ротовому аппарату наутилуса, состоящему из кальцита, укрепленного хитином. На дне океанов находят целые россыпи челюстей более поздних головоногих моллюсков, кое-где они покрывают дно сплошным слоем. Однако лишь в 1965 г. в такого рода осадках были достоверно определены аммониты — по необыкновенно хорошо сохранившемуся экземпляру челюстей, который нашли исследователи из Университета в Монтевидео. Когда раскололи найденный фосфатный желвак, внутри оказался ротовой аппарат моллюска: похожая на терку радула, лежащая между двумя уплощенными челюстями.

За 10 лет многое прояснилось, хотя нерешенные вопросы еще есть. Профессор Леман, выбрав подходящий экземпляр аммонита, изучил его внутреннюю структуру: он стачивал кальцит по десятой доле миллиметра, открывая и зарисовывая темные контуры в белой породе, заполняющей камеру, в которой помещалось тело моллюска. Леман разобрался в строении замыкательного аппарата. В знакомой, но не до конца понятной структуре он обнаружил пару пригнанных друг к другу лопастей — несомненно, это были челюсти, состоявшие из ороговикованного углистого вещества. По этому образцу была изготовлена объемная модель.

Когда исследовали «зубы» и челюсти, все предстало в новом свете. Тупоконечная нижняя челюсть аммонита даже отдаленно не напоминала режущее орудие или острый клюв. Челюстные мышцы не были мощными. Животное, видимо, медленно плавало вдоль дна, собирая челюстью, как совком, придонные организмы, которые и служили пищей. Эти данные неоднократно подтверждались. Так, Леман обнаружил в лондонских коллекциях невзрачный черный обломок породы, в котором оказалось несколько раковин размером чуть больше сантиметра. В некоторых раковинах сохранились челюсти, а в одной — даже остатки последней трапезы моллюска: мелкие фораминиферы и остракоды.

Разнообразие форм аммонитов — от привычных прихотливых спиралей до нетипичных форм, больше похожих на саксофон, чем на характерную «валторну» — теперь кажется понятным: форма раковины, вероятно, не имела приспособительного значения, поскольку для этих медленно дрейфующих у самого дна животных сопротивление воды не играло особой роли. Аммониты обитали в районах континентальных шельфов. Должно быть, они, как и современные морские улитки, эти санитары морского дна, подчистую подбирали всякую мелкую добычу, и живую и мертвую. В Южной Дакоте нашли аммонит, в раковине которого было 16 дырочек — расположение их точно совпадало с расположением зубов одной из современных ему морских рептилий. Находили и раковины, срезанные, как ножницами, клешнями каких-то крупных крабов. Безобидным подводным корабликам пришел конец, когда появились могучие костные рыбы и новые прогрессивные безраковинные головоногие. Может быть, аммониты вымерли не столько потому, что хищники пожирали взрослых моллюсков, сколько потому, что мелкие личиночные стадии гибли массами в ужесточившейся борьбе за существование

Модель раздувающейся Вселенной

СОГЛАСНО стандартной космологической модели, Вселенная начала расширяться из состояния бесконечно большой плотности и через 10^{-35} с ее объем достиг размеров бейсбольного мяча. На этой стадии она была уже слишком велика и расширялась слишком быстро, чтобы сигнал, распространяющийся со скоростью света, мог ее пересечь. Причинная связь между событиями, определяемая скоростью света, была ограничена областью диаметром 10^{-25} см, т.е. во Вселенной размером с бейсбольный мяч поместилось бы 10^{78} таких областей.

Современная теоретическая физика позволяет решить несколько сложнейших космологических проблем в предположении, что наблюдаемая Вселенная когда-то была чрезвычайно малой частью такой причинно-связанной области. В период от 10^{-35}с после начала расширения вплоть до 10^{-32}с эта область могла пройти через так называемую «инфляционную» фазу. При этом ее размер увеличился в 10^{100} раз — от триллионной доли диаметра протона до величины, в 10^{47} раз превышающей размер наблюдаемой Вселенной. (Точное значение «инфляционного» коэффициента несущественно, если он больше 10^{25}.) К концу этого грандиозного расширения область, которой предстояло стать современной наблюдаемой Вселенной, должна была достичь размера бейсбольного мяча, вкрапленного во Вселенную огромных размеров, которая продолжала расширяться за пределы современного горизонта. Спустя $10^{-35} \div 10^{-32}$с после начала расширения, «инфляционная» фаза сменилась фазой обычного расширения, и область размером с бейсбольный мяч, так же как и окружающая Вселенная, обрела скорость расширения, которая с тех пор и обусловливает ее эволюцию.

Эта удивительная гипотеза была выдвинута группой физиков и космологов, которые обратились к физике элементарных частиц, чтобы построить последовательную картину Вселенной в возрасте 10^{-35} с. Впервые эта идея была изложена А. Гутом (Массачусетский технологический институт) в статье, опубликованной в 1982 г. в журнале "Physical Review". Гут указал на некоторые противоречия модели, однако А. Олбрехту и П. Стейнхардту (Пенсильванский университет) и А. Линде (ФИАН, СССР) удалось преодолеть ряд трудностей. В настоящее время они и несколько других исследователей пересматривают модель раздувающейся Вселенной и изучают ее следствия.

Основная идея Гута заключалась в том, чтобы применить принципы термодинамики к веществу и энергии во Вселенной с температурой около 10^{27}K. При такой колоссальной температуре частицы имеют энергию, при которой, по-видимому, перестает существовать различие между тремя основными фундаментальными взаимодействиями. Для обычных температур характерны четыре типа фундаментальных взаимодействий: сильное взаимодействие, благодаря которому существуют ядра атомов, слабое взаимодействие, ответственное, например, за β-распад нейтрона, а также электромагнитное и гравитационное взаимодействия. Заветная мечта физиков-теоретиков — дать единое объяснение всем четырем типам фундаментальных взаимодействий, но к настоящему времени удалось связать лишь первые три в рамках теории великого объединения. Эта теория предсказывает, что при температуре выше критического значения, 10^{27}K, сильное, слабое и электромагнитное взаимодействия симметричны в том смысле, что они неотличимы друг от друга. При температуре ниже критической сильное взаимодействие отличается от двух других и симметрия нарушается. По предположению Гута, одно из следствий нарушения симметрии — переход вещества и энергии в новую фазу.

Можно провести аналогию между фазовым переходом, предсказанным Гутом, и переходом вещества из жидкого состояния в твердое. В последнем случае фазовый переход сопровождается образованием кристаллической решетки. В результате такой перестройки полная энергия системы снижается и выделяется некоторое количество энергии, называемое скрытой теплотой. Однако если в жидкости нет центров кристаллизации, то ее можно охладить ниже точки замерзания. Эту новую фазу называют переохлажденной жидкостью. Если поместить в такую жидкость подходящие ядра кристаллизации, то она замерзнет. При этом выделится скрытая теплота и температура твердого вещества повысится.

По аналогии Гут установил, что в течение симметричной фазы, предсказываемой теорией великого объединения, на ранних стадиях эволюции Вселенной могло наступить переохлаждение. Такая симметричная фаза может возникнуть в некоторых областях Вселенной при температуре всего 10^{22}K. В переохлажденной симметричной фазе расширение происходит экспоненциально, до тех пор пока не будет нарушена симметрия между фундаментальными взаимодействиями и наступит фаза, следы которой теперь проявляются всюду в наблюдаемой Вселенной. Такой переход в расширяющейся области, находящейся в симметричной фазе, сопровождается выделением огромного количества скрытой теплоты, и область вновь достигает температуры около 10^{27}K.

Хотя в модели раздувающейся Вселенной есть еще некоторые нерешенные проблемы, она позволяет объяснить по крайней мере три загадки космологии. Одна из них — однородность реликтового (микроволнового фонового) излучения, которая наблюдается с точностью до 1/10000. Если стандартная космологическая модель верна, то области неба, угловое расстояние между которыми превосходит 2°, не должны были взаимодействовать в течение эволюции Вселенной. Однородность реликтового излучения нельзя объяснить взаимодействием между удаленными областями, поэтому в стандартную космологическую модель приходится вводить гипотезы *ad hoc*. В модели раздувающейся Вселенной этот факт легко объясняется. Ведь «предтеча» современной наблюдаемой Вселенной значительно меньше, чем предсказывается в рамках стандартной космологической модели, поэтому когда-то каждая часть наблюдаемой Вселенной была причинно связана с любой другой ее частью.

Вторая загадка — почему трехмерное пространство можно считать евклидовым? Основной вопрос, на который должна дать ответ космология: открыта или замкнута Вселенная, т.е. будет ли расширение продолжаться вечно или средняя плотность вещества достаточна для того, чтобы под действием силы гравитации расширение прекратилось и сменилось сжатием? Если средняя плотность вещества точно равна критическому значению, то расширение никогда не прекратится. Такая Вселенная называется плоской — пространство не искривлено и справедлива евклидова геометрия.

Средняя плотность вещества во Вселенной до сих пор точно не определена, но известно, что она лежит в пределах плюс-минус один порядок величины от критического значения. Для более глубокого понимания ранних стадий эволюции Вселенной важнее знать не то, открыта она или замкнута, а то, как средняя плотность вещества в наблюдаемой Вселенной достигла значения, близкого к критическому. Согласно стандартной космологической модели, плотность ранней Вселенной должна была равняться критической с точностью до 10^{-49}, чтобы в ходе последующей эволюции был достигнут современный диапазон значений. В модели раздувающейся Вселенной такая

точность не требуется. В ходе экспоненциального расширения область, занятая наблюдаемой Вселенной, становится все более плоской и средняя плотность вещества естественным образом достигает критического значения. Наблюдаемая Вселенная — плоская по той же причине, по которой можно считать плоскостью небольшой участок поверхности шара.

Третья загадка связана с образованием «экзотических» частиц — так называемых магнитных монополей, предсказываемых теорией великого объединения. Магнитные монополи рождаются в течение перехода от симметричной фазы к фазе, в которой симметрия между фундаментальными взаимодействиями нарушена. В рамках стандартной космологической модели такой фазовый переход происходит настолько быстро, что должно было появиться огромное число магнитных монополей, способное сильно увеличить среднюю плотность вещества. В модели раздувающейся Вселенной фазовый переход происходит значительно медленнее, и поэтому возникает гораздо меньше монополей.

Правда, модель раздувающейся Вселенной не способна объяснить важнейшую наблюдаемую особенность современного мира: скучивание вещества, проявляющееся в существовании звезд и галактик. Небольшие флуктуации плотности в ранней Вселенной должны были либо полностью рассеяться, либо слишком быстро сжаться. По мнению некоторых исследователей, скопления галактик, возможно, образовались значительно позже.

Автокаталитический сплайсинг

ПОЧТИ во всех живых организмах генетическая информация закодирована в ДНК, а посредством транскрипции переносится в РНК трех различных типов: матричную РНК, на которой при трансляции образуются белки, рибосомную РНК и транспортную РНК, которые играют важную роль в трансляции. Различные стадии транскрипции и трансляции катализируются ферментами. Например, нуклеазы расщепляют нуклеиновые кислоты на фрагменты, лигазы «сшивают» такие фрагменты друг с другом или соединяют их в кольца. Считалось само собой разумеющимся, что для каждой такой реакции есть свой фермент и что нуклеиновые кислоты не способны расщепляться на фрагменты и вновь соединяться самостоятельно. Но так ли это? Совсем недавно сотрудники Колорадского универститета (г. Боулдер) получили веские доказательства в пользу того, что у одноклеточного организма *Tetrahymena thermophila* есть РНК, которая сама осуществляет как раз такую реакцию.

У некоторых видов *Tetrahymena* гены, кодирующие рибосомную РНК, как бы разорваны, т.е. кодирующая часть генов прерывается так называемыми вкрапленными последовательностями, или интронами. Ген транскрибируется целиком с образованием предшественника рибосомной РНК, а затем происходит его сплайсинг: интрон вырезается из предшественника, а две длинные последовательности, которые располагались до и после интрона, соединяются вместе ковалентной связью с образованием зрелой рибосомной РНК. Пока ничего удивительного: еще в 1977 г. было известно, что у организмов, стоящих на эволюционной лестнице выше бактерий, во многие гены матричных, рибосомных и транспортных РНК вкраплены интроны. Однако, исследуя сплайсинг у *T. thermophila*, Т. Цек и его коллеги из Колорадского университета столкнулись с неожиданностями. Во-первых, оказалось, что интрон, вырезанный из РНК-предшественника в виде линейной молекулы, вскоре замыкается в кольцо. Во-вторых, анализируя последовательность РНК на ее начальном конце (т.е. на 5′-конце) обнаружили лишний гуанозиновый нуклеотид — лишний в том смысле, что он не закодирован в ДНК, с которой считывается рибосомная РНК.

Но совсем уже странные вещи начались, когда Цек и его сотрудники попытались детально исследовать транскрипцию ДНК с образованием РНК: оказалось, что предшественник РНК короткоживущий. Даже в искусственной системе, в которой сплайсинг не должен происходить, интрон быстро вырезался из предшественника и замыкался в кольцо. Никакие белки в этом процессе, судя по всему, не участвовали. Варьируя концентрацию соли в системе, исследователи сумели разобщить транскрипцию и сплайсинг, так что предшественник РНК накапливался. А когда к нему добавляли достаточное количество гуанозинтрифосфата (GTP), начинались сплайсинг и образование кольцевых молекул, причем даже если предварительно все белки были тщательно удалены. Было показано, что GTP нужен на стадии вырезания: одна молекула GTP внедряется между 5′-концевым экзоном и началом интрона, другая встраивается у 3′-конца интрона, а затем вклинивается у 5′-конца, и в результате интрон замыкается в кольцо.

И вот теперь те же исследователи из Колорадского университета получили более строгие доказательства существования автосплайсинга и самозамыкания в кольцо. Результаты своих экспериментов К. Крюгер, П. Грабовски, А. Зауг, Дж. Сэндз, Д. Готтшлинг и Т. Цек опубликовали в журнале "Cell". Они встроили интрон и примыкающие к нему участки экзонов в плазмиду (плазмида — это небольшая кольцевая молекула ДНК из бактерий), затем очищенную плазмидную ДНК транскрибировали с помощью бактериальной полимеразы, которая на ДНК-матрице соединяет отдельные нуклеотиды в молекулу РНК. Полимеразу удаляли, так что в системе не оставалось белков. И вновь интрон сам вырезался и замыкался в кольцо. Цеку и его сотрудникам удалось также продемонстрировать, что в ходе той же реакции, в которой вырезается интрон, сшиваются два экзона. Есть основания считать, что во многих отношениях интрон *T. thermophila* действует как фермент. Цек назвал его рибозимом.

Распространены ли в природе рибозимы? Этого пока никто не знает. Пока что идентифицировано мало ферментов сплайсинга, и вполне может оказаться, что автокатализ — широко распространенное явление (хотя маловероятно, что он имеет место при сплайсинге предшественников матричных РНК, для которого, по-видимому, не требуется целостности последовательности интронов). Впрочем, достаточно оказалось обнаружить одинединственный случай автосплайсинга, чтобы появились новые идеи о функциях РНК. В отличие от двухспиральной ДНК одноцепочечная РНК легко сворачивается в структуры самой причудливой формы, а ведь типичным ферментам как раз свойственна сложная специфическая трехмерная конфигурация молекулы. Последнее время все чаще говорят о том, что некоторые РНК могут играть в трансляции и других процессах не только чисто структурную, но и активную роль. Помимо транспортных и рибосомных РНК, в клетке имеется много различных комплексов небольших молекул РНК с белками; функции этих комплексов мы только начинаем понимать. Многие исследователи, задумывающиеся о происхождении жизни, склоняются к тому, что первичным генетическим материалом была именно РНК, что она предшествовала ДНК. Благодаря способности примитивных РНК к автокатализу исходный генетический материал мог варьировать, обеспечивая тем самым изменчивость — главное условие эволюции.

Библиография

МАГИЯ МАТЕМАТИКИ

LISP 1.5 PROGRAMMER'S MANUAL. John McCarthy. The MIT Press, 1962.

THE LITTLE LISPer. Daniel P. Friedman. Science Research Associates, Inc., 1974.

LET'S TALK LISP. L.S. Siklóssy. Prentice-Hall, Inc., 1976.

LISP. Patrick Henry Winston and Berthold K. Horn. Addison-Wesley Publishing Co., Inc., 1981.

БУДУЩЕЕ СЕЛЬСКОГО ХОЗЯЙСТВА США

FARM AND FOOD POLICY: ISSUES OF THE 1980s. Don Paarlberg. University of Nebraska Press, 1980.

THE FUTURE OF AMERICAN AGRICULTURE AS A STRATEGIC RESOURCE. Edited by Sandra S. Batie and Robert G. Healy. The Conservation Foundation, 1980.

Вольф М. Б., Дмитриевский Д. Ю. География мирового сельского хозяйства. — М.: Мысль, 1981.

Меньшиков М. А. Американский путь развития сельского хозяйства и его международные аспекты. — США: экономика, политика, идеология, 1981, № 1.

Продовольственная проблема в современном мире. — М.: Наука, 1983. Экономика США, под. ред. В. Ф. Фелезовой. — М.: изд-во МГУ, 1983.

РЕШЕТОЧНАЯ ТЕОРИЯ УДЕРЖАНИЯ КВАРКОВ

THE CONFINEMENT OF QUARKS. Yoichiro Nambu in *Scientific American*. Vol. 235, No. 5, pages 48 — 60; November, 1976.

THE BAG MODEL OF QUARK CONFINEMENT. Kenneth A. Johnson in *Scientific American*, Vol. 241, No. 1, pages 100 — 109; July, 1979.

GAUGE THEORIES OF THE FORCES BETWEEN ELEMENTARY PARTICLES. Gerard't Hooft in *Scientific American*, Vol. 242, No. 6, pages 90 — 116, June, 1980.

LATTICE GAUGE THEORIES AND MONTE CARLO SIMULATIONS. Claudio Rebbi in *Non-Perturbative Aspects of Quantum Field Theory*. World Scientific Publishing Co., Singapore, 1982.

Шехтер В. М. Кварки. — Природа, 1970, № 2, с. 53 — 59.

СИНТЕТИЧЕСКИЕ ВАКЦИНЫ

ANTIBODIES TO CHEMICALLY SYNTHESIZED PEPTIDES PREDICTED FROM DNA SEQUENCES AS PROBES OF GENE EXPRESSION. Richard A. Lerner, J. Gregor Sutcliffe and Thomas M. Shinnick in *Cell*, Vol. 23, No. 2, pages 309 — 310; February, 1981.

TAPPING THE IMMUNOLOGICAL REPERTOIRE TO PRODUCE ANTIBODIES OF PREDETERMINED SPECIFICITY. Richard A. Lerner in *Nature*, Vol. 299, No. 5884, pages 592 — 596; October, 14, 1982.

ОПТИЧЕСКИЕ КОМПЬЮТЕРЫ

OPTICAL BISTABILITY. Edited by Charles M. Bowden, M. Ciftan and H. R. Robl. Plenum Press, 1981.

DYNAMIC NON-LINEAR OPTICAL PROCESSES IN SEMICONDUCTORS. Alan Miller, David A. B. Miller and S. Desmond Smith in *Advances in Physics*, Vol. 30, No. 6, pages 697 — 800; November/December, 1981.

BISTABLE OPTICAL DEVICES: PHYSICS AND OPERATING CHARACTERISTICS. David A. B. Miller in *Laser Focus with Fiberoptic Technology*, Vol. 18, No. 4, pages 79 — 84; April, 1982.

СКРЫТЫЕ ЗРИТЕЛЬНЫЕ ПРОЦЕССЫ

EYE TORSION AND VISUAL TILT ARE MEDIATED BY DIFFERENT BINOCULAR PROCESSES. Jeremy M. Wolfe and Richard Held in *Vision Research*, Vol. 19, No. 8, pages 917 — 920; 1979.

CYCLOPEAN STIMULATION CAN INFLUENCE SENSATIONS OF SELF-MOTION IN NORMAL AND STEREOBLIND SUBJECTS. Jeremy M. Wolfe and Richard Held in *Perception and Psychophysics*, Vol. 28, No. 2, pages 139 — 142; 1980.

INTEROCULAR TRANSFER OF VISUAL AFTEREFFECTS. R. Blake, R. Overton and S. Lema-Stern in *Journal of Experimental Psychology: HPP*, Vol. 7, pages 367 — 381; 1981.

IS ACCOMMODATION COLORBLIND? FOCUSING CHROMATIC CONTOURS. Jeremy M. Wolfe and D. Alfred Owens in *Perception*, Vol. 10, pages 53 — 62; 1981.

A PURELY BINOCULAR MECHANISM IN HUMAN VISION. Jeremy M. Wolfe and Richard Held in *Vision Research*, Vol. 21, No. 12, pages 1755 — 1759; 1981.

АКТИВНАЯ СОЛНЕЧНАЯ КОРОНА

THE SOLAR CORONA. Jay M. Pasachoff in *Scientific American*, Vol. 229, No. 4, pages 68 — 79; October, 1973.

A NEW SUN: THE SOLAR RESULTS FROM SKYLAB. John A. Eddy. Scientific and Technical Information Office, National Aeronautics and Space Administration, U.S. Government Printing Office, 1979.

ЯДОВИТЫЕ ЛЯГУШКИ

PRELIMINARY EVALUATION OF SKIN TOXINS AND VOCALIZATIONS IN TAXONOMIC AND EVOLUTIONARY STUDIES OF POISON-DART FROGS (DENDROBATIDAE). Charles W. Myers and John W. Daly in *Bulletin of the American Museum of Natural History*, Vol. 157, Article 3, pages 173 — 262; 1976.

MORPHOLOGY OF THE GRANULAR SECRETORY GLANDS IN SKIN OF POISON-DART FROGS (DENDROBATIDAE). Maria Neuwirth, John W. Daly, Charles W. Myers and Lois W. Tice in *Tissue and Cell*, Vol. 11, No. 4, pages 755 — 771; 1979.

TAXONOMY AND ECOLOGY OF *DENDROBATES BOMBETES*. A NEW ANDEAN POISON FROG WITH NEW SKIN TOXINS. Charles W. Myers and John W. Daly in *American Museum Novitates*, No. 2692, pages 1 — 23; January 29, 1980.

THE PHARMACOLOGY OF PUMILIOTOXIN-B. I: INTERACTION WITH CALCIUM SITES IN THE SARCOPLASMIC RETICULUM OF SKELETAL MUSCLE. E. X. Albuquerque, J.E. Warnick, M.A. Maleque, F.C. Kauffman, R. Tamburini, Y. Nimit and J.W. Daly in *Molecular Pharmacology*, Vol. 19, No. 3, pages 411 — 424; May, 1981.

ВЕЕРНЫЕ СВОДЫ

MEDIEVAL STRUCTURE: THE GOTHIC VAULT. James J. Acland. University of Toronto Press, 1972.

FAN VAULTING: A STUDY OF FORM, TECHNOLOGY AND MEANING. Walter C. Leedy, Jr. Arts and Architecture Press, Santa Monica, 1980.

EXPERIMENTS IN GOTHIC STRUCTURE. Robert Mark. The MIT Press, 1982.

НАУКА ВОКРУГ НАС

OBSERVATION OF OSCILLATORY ICE STATES IN A SIMPLE CONVECTION EXPERIMENT. Pierre Welander in *Journal of Geophysical Research*, Vol. 82, No. 18, pages 2591 — 2592; June 20, 1977.

В МИРЕ НАУКИ

Подписано в печать 06.04.83.
По оригинал-макету. Формат 60 × 90/8
Гарнитуры таймс, гелиос
Офсетная печать.
Объем 6,75 бум. л.
Усл.-печ. л. 13,50,
Печ. л. 6,75, Уч.-изд. л. 17,85.
Усл. кр.-отт. 46,86.
Изд. № 36/2664. Заказ 203-83.
Тираж 20000 экз. Цена 2 р.
Издательство «Мир»
Набрано в редакции по подготовке
оригинал-макетов издательства «Мир»
на фотонаборном комплексе
«Компьюграфик»
Типография В/О Внешторгиздат
Главиздатэкспорта Госкомиздата СССР
127349, Москва, И-349,
Илимская, 7

Издательство **МИР** *предлагает:*

МЕХАНИКА ОБРАЗОВАНИЯ ВОРОНОК ПРИ УДАРЕ И ВЗРЫВЕ

Перевод с английского



В книге рассматривается механизм образования воронок на поверхности грунта в результате падения метеорита или взрыва. Анализ ударного или взрывного воздействия и последующего состояния кратеров проводится с учетом влияния прочности и неоднородности грунта, силы тяжести и атмосферного давления. Выяснено сходство и различие процессов при ударе и взрыве. Широко используются результаты исследований кратеров на лунной поверхности. Показана эффективность применения голографии для наблюдения за поверхностными волнами в грунте. Среди авторов статей — известные специалисты США по механике ударных и взрывных процессов и космическим исследованиям.

Книга будет полезна широкому кругу специалистов, работающих в области механики взрывных процессов, космических исследований и механике грунтов.

1977, 228 с.

Цена 1 р. 48 к.

В следующем номере:



ПАМЯТЬ У ПТИЦ — НОВЫЕ ИССЛЕДОВАНИЯ

ХИМИЧЕСКИЕ ПРОЦЕССЫ В ПЛАМЕНИ

МИКРОПРОГРАММИРОВАНИЕ — СРЕДСТВО РАСШИРЕНИЯ ВОЗМОЖНОСТЕЙ ЭВМ

ПИСЬМЕННОСТЬ ДРЕВНЕЙШЕЙ ЦИВИЛИЗАЦИИ ДОЛИНЫ ИНДА

МИТОХОНДРИАЛЬНАЯ ДНК

О БУДУЩЕМ ВСЕЛЕННОЙ

КОЛЕБАТЕЛЬНЫЕ ХИМИЧЕСКИЕ РЕАКЦИИ

АУТОПСИЯ — ЕЕ РОЛЬ В ПЛАНИРОВАНИИ ЗДРАВООХРАНЕНИЯ

ЗАДАЧА АВТОЛЮБИТЕЛЮ:
КАК ПРОАНАЛИЗИРОВАТЬ РАБОТУ
СИСТЕМЫ ПЕРЕКЛЮЧЕНИЯ СВЕТОФОРОВ
НА АВТОТРАССЕ

РЕКУРСИВНЫЙ ХОРОВОД
В ЯЗЫКЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ЛИСП

ИНДЕКС 91310 ISSN 0208-0621